

A MODEL OF THE SHORT-CHANNEL,
METAL-OXIDE-SEMICONDUCTOR FIELD-EFFECT TRANSISTOR
FOR PRAGMATIC MIXED-MODE DEVICE/CIRCUIT SIMULATION

By

KEITH R. GREEN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1993

ACKNOWLEDGEMENTS

I am grateful for the opportunity to have worked with Professor Jerry Fossum who provided me invaluable technical training, and while doing so made my graduate-school experience exciting and enjoyable.

I thank Professor Mark Law for his guidance in matters of fabrication technology. I also appreciate Professors Burk, Fox and Mataga, for their willingness to serve on my supervisory committee.

I thank Ian Young for his technical guidance and for supplying device measurement data which made an an important contribution to this dissertation.

I also thank Minchang Liang for help with PISCES simulations of bipolar junction transistors, and Srinath Krishnan for many useful discussions about semiconductor device physics.

Finally, I am especially grateful for the emotional support provided by my family during the course of this work.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS.....	ii
ABSTRACT.....	v
CHAPTERS	
1 INTRODUCTION.....	1
2 A PRAGMATIC APPROACH TO INTEGRATED PROCESS/DEVICE/CIRCUIT SIMULATION FOR IC TECHNOLOGY DEVELOPMENT.....	8
2.1 Introduction.....	8
2.2 SUPREM-3/SUMM/MMSPICE.....	9
2.3 Verification/Examples.....	14
2.4 Conclusions.....	30
3 A MODEL OF THE SHORT-CHANNEL MOSFET FOR PRAGMATIC MIXED-MODE DEVICE/CIRCUIT SIMULATION.....	32
3.1 Introduction.....	32
3.2 Determining the Region of Operation.....	35
3.2.1 Charge Sharing.....	39
3.2.2 DIBL and DICE.....	42
3.3 Strong Inversion Channel Transport Current.....	47
3.3.1 Triode Current.....	49
3.3.2 Saturation Current.....	51
3.4 Weak Inversion Channel Transport Current.....	55
3.5 Cubic Spline for Moderate Inversion Operation...	67
3.6 Charging Currents of the Intrinsic Device.....	68
3.6.1 Gate Charge.....	71
3.6.2 Drain Charge.....	72
3.6.3 Source Charge.....	73
3.6.4 Body Charge.....	73
3.7 Effects of LDD/LDS.....	73
3.7.1 Voltage Drop Across the LDS.....	74
3.7.2 Voltage Drop Across the LDD.....	74
3.8 Generation Current Due to Impact Ionization.....	78
3.9 Summary.....	79

4 MODEL IMPLEMENTATION IN SUMM/MMSPICE.....	80
4.1 Introduction.....	80
4.2 Description of Equivalent Circuit Elements.....	82
4.3 Implementation in MMSPICE.....	83
4.4 Device and Model Parameters.....	86
4.5 Model Parameter Evaluation in SUMM and MODCHK...	91
4.6 Implementation in SUMM.....	99
4.7 FISCES Support.....	105
5 MODEL VERIFICATION/APPLICATION.....	110
5.1 Introduction.....	110
5.2 Comparisons with Device and Circuit Measurements.....	110
5.3 Application to Technology Scaling.....	120
6 SUMMARY AND SUGGESTIONS FOR FUTURE WORK.....	127
APPENDIX	
EVALUATION OF THE MMSPICE BJT MODEL PARAMETERS JEO AND TE.....	130
REFERENCES.....	134
BIOGRAPHICAL SKETCH.....	138

Abstract of Dissertation Presented to the Graduate School of
the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

A MODEL OF THE SHORT-CHANNEL,
METAL-OXIDE-SEMICONDUCTOR FIELD-EFFECT TRANSISTOR
FOR PRAGMATIC MIXED-MODE DEVICE/CIRCUIT SIMULATION

By

KEITH R. GREEN

August 1993

Chairman: Dr. J. G. Fossum
Major Department: Electrical Engineering

This dissertation is concerned with physical modeling of scaled metal-oxide-semiconductor field-effect transistors (MOSFETs) and the development of a pragmatic mixed-mode device/circuit simulator. The simulator enables an efficient means of identifying and optimizing key short-channel MOSFET structural features that control device and circuit performances, and complements an analogous tool for bipolar junction transistors (BJTs). The device/circuit simulator relies on a new charge-based model for bulk-silicon surface-channel MOSFETs. Models for short-channel effects are derived from quasi-two-dimensional analyses of the complex intrinsic-region two-dimensional electric field prevalent in scaled (submicron) MOSFET technologies; these analyses are linked with one of the electric field distribution in the lightly-doped drain region. Iterative solution techniques are required to characterize transport current, impact-ionization

current, and quasi-static regional charges (for dynamic charging currents). The MOSFET model is implemented in the circuit simulator SPICE2 to create the basis (MMSPICE) of the mixed-mode tool. The unique evaluation of the physical (structure-dependent) device parameters ensures proper parametric correlations for accurate predictive simulation, and facilitates integration with a process simulator. The parameter-evaluation routines are incorporated in a menu-driven computer program (SUMM) that accepts as input structural information from a process simulator, and automatically assembles the input file for the device/circuit simulator. Device and circuit simulations are compared with measurements and rigorous numerical simulations to provide support for the new model and mixed-mode tool. Application of the tool to scaling a merged MOSFET/BJT integrated-circuit technology is described.

CHAPTER 1 INTRODUCTION

Technology CAD (TCAD), which requires integrated, physics-based, CPU-efficient tools for predictive process, device, and circuit simulation, is now a crucial part of integrated-circuit (IC) manufacturing [Llo90]. Although useful tools exist, they are not well integrated, and they generally lack adequate physical basis and/or acceptable computational efficiency. For example, in the conventional IC TCAD system (Fig. 1.1(a)), a robust numerical device simulator, driven by a numerical process simulator, is used to optimize (e.g., via least-squares fitting) empirical device model parameters for circuit simulation. The needed optimization misses correlations between device model and technology parameters. Hence, the conventional integrated system is not truly physical, and could yield nonunique (erroneous) circuit simulations for a specific process flow. Furthermore, the conventional system, based on the presumption of exactness, tends to be inefficient since its numerical rigor is incompatible with inherent uncertainties in the physical modeling underlying the process and device simulators.

Because of the tool-integration problem, there is interest in "mixed-mode" device/circuit simulation, in which the coupled circuit nodal equations and carrier transport

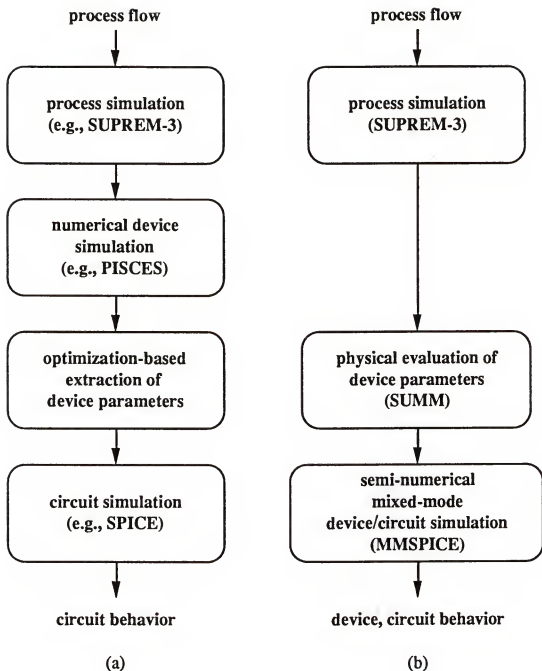


Fig. 1.1. Flowchart of (a) conventional and (b) application-specific technology-CAD systems.

equations are solved simultaneously. However purely numerical mixed-mode simulators exacerbate the computational inefficiency mentioned above, and indeed they commonly have convergence problems.

Recently the use of application-specific analytic TCAD tools was suggested as a viable way of enhancing computational efficiency [Mar88]. Exploiting this approach, we suggest a pragmatic alternative to conventional TCAD based on the incorporation of application-specified (i.e., structure-specific) analytic, but seminumerical, device models in the circuit simulator. These models are characterized by physical parameters that relate directly to the device structure, and therefore are actually device simulators. Hence, the integration of the device and circuit simulators is natural in this mixed-mode tool, which can be easily integrated with the process simulator by a program that evaluates the model parameters from the doping profile (Fig. 1.1(b)). Unlike conventional optimization-based parameter extraction, the evaluation of these parameters is physical, based on an analysis of the device structure with criteria consistent with the device physics. Therefore, all parametric correlations are inherently accounted for, and the real sensitivities of device and circuit performances to variations in the fabrication process can be reliably and efficiently predicted.

A bipolar junction transistor (BJT) model that meets the criteria required for integration into the pragmatic TCAD

system was developed previously [Jeo89]. That model is based on linked regional analyses of the device structure, and provides a physical alternative to the widely used empirical Gummel-Poon/SPICE model [Get78] for advanced BJTs [Zde87]. For example, it contains a comprehensive accounting of high collector-current effects (quasi-saturation); and, unlike the Gummel-Poon/SPICE model which relies on empirical transit times to characterize the majority-carrier integral in the integral charge-control relation, the new model relies on an analytic formulation of that integral in terms of technology parameters. With its implementation in SPICE2 [Nag75] and incorporation of the parameter evaluator SUMM [Gre90], a viable mixed-mode BJT device/circuit simulator was created.

The deficiencies of existing short-channel MOSFET circuit models necessitate the development of a new predictive model for this pragmatic system. As MOSFET gate lengths were scaled to submicron dimensions, short-channel circuit models were developed via empirical extensions of well characterized long-channel models. In some models these extensions arose from empirical modifications to simple one-dimensional (1-D) characterizations of electric field to account for short-channel effects that manifest from complex two-dimensional (2-D) electric field distributions; an example is the 1-D weak-inversion channel-length modulation model of Chan et al. [Cha87]. Other MOSFET models do not account explicitly for the various short-channel effects; for example, the familiar BSIM model [She87] relies heavily on

empirical "fitting" expressions that are intended to account for these effects without distinguishing them. Such approaches were taken to preserve the accuracy and efficiency required for simulation of large circuits. As a consequence though, MOSFET circuit models tend to be deficient in their ability to predict short-channel behavior, and they provide little insight into how short-channel effects are related. In order to achieve predictive simulation, analyses of the 2-D electric field is necessary. However, rigorous 2-D analyses require computational expense that is undesirable in circuit simulation; therefore an alternative modeling approach that relies on approximate solutions to the 2-D electric field problems is preferable. This dissertation presents a model that relies on such quasi-2-D analyses to account for short-channel effects, viz., drain-induced barrier lowering [Tro79], drain-induced conductivity enhancement [Vee88], charge-sharing and channel-length modulation [Sze81]. These solutions of quasi-2-D analyses are physically linked, and are expressed in simple terms that yield physical insight into device behavior as well as a predictive circuit model, which also has utility in TCAD, e.g., for statistical simulation.

Chapter 2 exemplifies the pragmatic TCAD system depicted in Fig. 1.1(b) for advanced bipolar technologies. This simulation system of SUPREM-3/SUMM/MMSPICE is verified with measurements and compared with the conventional system of Fig. 1.1(a).

Chapter 3 presents the new physical short-channel MOSFET circuit model that is based in part on quasi-2-D analyses. The model comprehensively accounts for short-channel effects, field-dependent mobility, lightly doped drain and source regions, impact ionization, and device parasitics. Iterative solution techniques are required in the model formalism to predict transport current, impact-ionization current, and quasi-static regional charges (to account for dynamic charging currents) from technology parameters with minimal reliance on empirical modeling.

Chapter 4 details the implementation of the new model in MMSPICE and the extension of the program SUMM for evaluating key MOSFET model parameters from a 1-D doping profile under the gate. This chapter also presents some support for the SUMM/MMSPICE tool provided by fully numerical device simulations.

Chapter 5 presents experimental support for the MOSFET model and mixed-mode device/circuit simulator. It does so by describing an application of the system to scaling a merged MOSFET/BJT IC technology.

Chapter 6 summarizes this work and presents suggestions for future related research.

An appendix details new evaluation algorithms for two MMSPICE bipolar model [Jeo89] parameters that characterize minority-carrier transport in the emitter region. These evaluations depend on the emitter surface recombination

velocity, which for polysilicon-contacted BJTs is critical in determining device current gain and operating speed.

CHAPTER 2
A PRAGMATIC APPROACH TO INTEGRATED PROCESS/DEVICE/CIRCUIT
SIMULATION FOR IC TECHNOLOGY DEVELOPMENT

2.1 Introduction

This chapter describes the development of a pragmatic integrated process/device/circuit simulation system (Fig. 1.1(b)) for advanced silicon bipolar technologies. Section 2.2 presents the system of SUPREM-3/SUMM/MMSPICE which comprises physics-based tools, allowing predictive seminumerical mixed-mode device/circuit simulation in a well integrated environment for computationally efficient bipolar IC TCAD. In Section 2.3 we verify its predictive capability with measurements and numerical device simulations using an enhanced version [Law90] of PISCES-II [Pin84]. Its utility is then demonstrated by simulations predicting how ECL propagation delay is sensitive to variations in an advanced bipolar fabrication process. An example shows the effect of varying the base (boron) ion-implantation dose on the transient characteristics of a three-stage (12-transistor) ECL circuit. Another example demonstrates the influence of the effective recombination velocity at the polysilicon-silicon interface of the BJT emitter on ECL circuit transients.

2.2 SUPREM-3/SUMM/MMSPICE

The SUPREM-3/SUMM/MMSPICE application-specific system comprises SUPREM-3 linked to MMSPICE by SUMM. MMSPICE [Jeo90] is the mixed-mode device/circuit simulator developed by writing a one-dimensional seminumerical bipolar junction transistor (BJT) model [Jeo89] into SPICE2 [Nag75] source code. The charge-based model physically characterizes all current components and regional charges in terms of technological parameters that are uniquely defined by the intrinsic doping-density profile of the advanced BJT structure. SUMM [Gre90] is a UNIX-based, menu-driven program that evaluates the model parameters from the tabulated one-dimensional doping profile such as that predicted by SUPREM-3 [Tec88], and automatically writes the model card in the circuit file for MMSPICE. Of the 31 model parameters, 12 are physical, including, for example, carrier saturated drift velocity. These are assigned typical, representative values. The other 19 parameters are structure-dependent, and are evaluated from the doping-density profile by numerical computation. The evaluation of these parameters ranges from direct evaluation, e.g., of metallurgical base width, to use of well-documented physical models, e.g., for carrier mobility-doping dependences used in the evaluation of the majority-carrier mobility in the epitaxial collector and the minority-carrier diffusivity in the base. The evaluations involve numerical analyses of Poisson's equation in the

junction regions to determine space-charge region boundaries and built-in potentials.

Evaluation criteria [Gre90] are defined to ensure the physical integrity of the model parameters, and hence to guarantee proper accounting for all parametric correlations. For example, the npn BJT model assumes an exponential net base doping profile expressed as

$$N_A^{\text{model}}(x) = NA0 \exp\left(-\frac{ETA}{WBM} x\right), \quad (2.1)$$

where model parameters are capitalized. WBM is the metallurgical base width, which is taken directly from the tabulated profile; NA0 is the extrapolated peak doping density at the emitter-base metallurgical junction; and ETA is the doping density gradient factor (see Fig. 2.1). The parameter ETA is evaluated under the imposed criterion of preserving the built-in electric field in the quasi-neutral base (QNB) caused by the doping-density gradient. This electric field is determined from the drift-diffusion balance of holes in equilibrium. For the model profile, with (2.1), the field is expressed as

$$E_{\text{model}} = \frac{kT}{q} \frac{d}{dx} [\ln N_A^{\text{model}}(x)] = -\frac{kT}{q} \frac{ETA}{WBM}. \quad (2.2)$$

The built-in electric field in the actual base profile ($N_A(x)$) is computed using

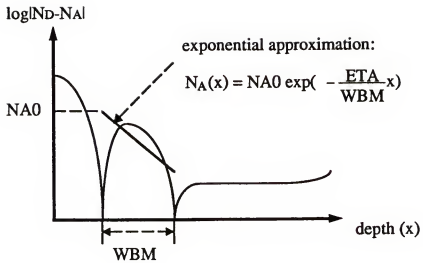


Fig. 2.1. BJT doping profile with exponential approximation for the base region superimposed.

$$E_{\text{actual}} = \frac{kT}{q} \frac{d}{dx} [\ln N_A(x)] \cong \frac{kT}{q} \left[\frac{\ln N_A(x_{c0}) - \ln N_A(x_{e0})}{x_{c0} - x_{e0}} \right] , \quad (2.3)$$

where we use an approximation based on the values of the actual doping density at the QNB boundaries, x_{e0} and x_{c0} , determined from numerical solutions of Poisson's equation in the junction regions [Gre90]. Imposing the stated criterion means equating (2.2) and (2.3), which leads to the evaluation of ETA:

$$\text{ETA} = -\text{WBM} \left[\frac{\ln N_A(x_{c0}) - \ln N_A(x_{e0})}{x_{c0} - x_{e0}} \right] . \quad (2.4)$$

With WBM and ETA, NAO is evaluated under the imposed criterion that the total net doping in the metallurgical base of the actual profile is preserved in the model:

$$\int_0^{\text{HEM}} N_A(x) dx = \int_0^{\text{HEM}} N_A^{\text{model}}(x) dx = \text{NAO} \frac{\text{WBM}}{\text{ETA}} [1 - \exp(-\text{ETA})] , \quad (2.5)$$

where the integral over the actual profile is evaluated numerically within SUMM. Thus NAO is evaluated directly via (2.5). Note that the criterion (2.5), with (2.4), virtually ensure equal (equilibrium) Gummel numbers from the actual profile and the model, and equal values for the equilibrium space-charge region depletion charges in the base at the emitter and collector junctions.

As another example of the seminumerical techniques used in SUMM to evaluate the doping-dependent BJT model

parameters, consider those related to the formation of the polysilicon emitter contact, which is most critical in advanced (scaled and self-aligned) bipolar technology. This formation is difficult to control and can yield substantive variations in the effective minority-hole surface recombination velocity, S_p [Eva90]. During the execution of SUMM, the user is prompted for the value of S_p (because it cannot be accurately predicted theoretically) to evaluate the model parameters TE , the minority-carrier charge-control time for the quasi-neutral emitter, and $JE0$, the saturation current density for the minority-carrier injection into the emitter. The evaluation of these two parameters, based on the the doping density $N_D(x)$ in the n^+ emitter, relies on an analytic model [Fos81] for the minority-hole transport. $JE0$, which depends on bandgap narrowing and other heavy-doping effects as well as on S_p , has a strong influence on the BJT current gain. TE governs the emitter delay of the BJT, and relates the hole charge in both the silicon and polysilicon parts of the emitter to the hole current.

The hole charge in the polysilicon, which is usually (as in PISCES-II [Pin84]) ignored without justification, is characterized in SUMM based on an assumed extended (silicon) emitter [Yun88]; i.e., there is no oxide at the polysilicon-silicon interface and grain boundaries in the polysilicon are insignificant. The neglect of any interfacial oxide is justified because it is avoided in typical advanced BJT technologies. Although the oxide tends to enhance the BJT

current gain in scaled technologies, it also tends to introduce processing instability and high emitter resistance [Eva90]. The formalism for evaluating JEO and TE is described in the appendix. It results in

$$JEO = \frac{qn_i^2}{N_{D(eff)}} \left(\frac{1}{S_p} + \frac{W_H}{D_{p0}} \right)^{-1} \left(1 + \frac{\tau_t}{\tau_A} \right) \quad (A1)$$

and

$$TE = \left(\frac{1}{\tau_t} + \frac{1}{\tau_A} \right)^{-1} + \frac{D_{p(poly)}}{S_p^2} \quad (A9)$$

for S_p high enough to imply no interfacial oxide. From (A1) and (A9) the correlation between TE and JEO, physically accounted for in SUMM, is apparent.

2.3 Verification/Examples

Verification of the integrated SUPREM-3/SUMM/MMSPICE system is provided by measurements and by purely numerical simulations using SUPREM-3 and PISCES-II. The results compared include both dc and ac BJT characteristics and their dependences on specific variations in the process flow of advanced bipolar technologies.

With SUPREM-3, the two npn polysilicon-contacted BJT doping profiles shown in Fig. 2.2 were generated. The profiles differ as a result of a factor-of-two variation in the base (boron) ion-implantation dose. Notice that in

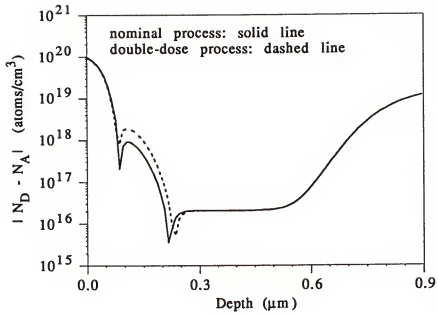


Fig. 2.2. SUPREM-3-simulated BJT doping-density profiles corresponding to two different (nominal and double) base-implant doses.

addition to altering the doping of the base, this specific process variation also affects the emitter and collector regions via compensation. Obviously then there are important correlations among the model parameters that are associated with the base, emitter, and collector regions. The parameter evaluation done in SUMM inherently accounts for these complex correlations e.g., how increasing the dose increased the metallurgical base width (WBM) and concomitantly decreased the epitaxial collector width (WEPI).

SUMM was used to evaluate model parameters for each profile, and subsequent MMSpice simulations yielded the Gummel plots ($V_{CE} = 2$ V) shown in Fig. 2.3 for an actual emitter area of $0.75 \mu\text{m} \times 4.0 \mu\text{m}$. Also shown in Fig. 2.3 are the Gummel plots predicted by PISCES [Law90]. Although there are discrepancies between the SUMM/MMSpice and PISCES predictions, the results do lend support to the former tool. Indeed the discrepancies can be attributed primarily to differences between the nonfundamental and/or ill-defined physical models assumed in the respective tools, which in fact render any device simulator somewhat equivocal. Two dimensional effects appear to be insignificant.

Both simulators predict negligible variation in the moderate- to high-bias base currents with the increase of implant dose. This is because the hole injection into the emitter (i.e., the predominant component of base current at these biases) is controlled by the heavily-doped emitter region, and therefore is not significantly affected by the

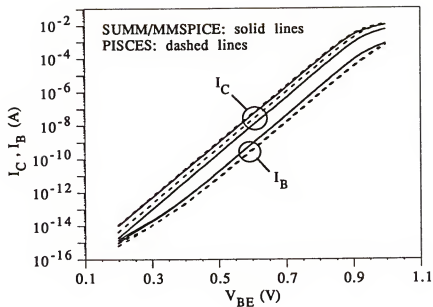


Fig. 2.3. Simulated Gummel plots ($V_{CE} = 2$ V) for the devices of Fig. 2.2. Note that doubling the base-implant dose decreases I_C , but affects I_B minimally.

increase in base doping. The hole injection depends on heavy-doping effects (e.g., bandgap narrowing) in the emitter. Differences in the modeling of these ill-defined effects account for discrepancies between the base currents predicted by the two simulators. Both simulators predict that base current arising from emitter-base space-charge region recombination becomes significant at low bias, with minor discrepancies caused by differences in the modeling of nonfundamental carrier lifetimes and of the ill-defined hole injection into the emitter. Note the predicted reduction of low-bias base current with the increase in base-implant dose, which is due to reduced emitter-base space-charge region recombination resulting from the modification of the junction doping.

The collector currents predicted by SUMM/MMSPICE and PISCES correspond well, showing discrepancies that can be attributed predominantly to differences in assumed physical models. For the nominal device at moderate bias, the SUMM/MMSPICE prediction is about 15% lower than that of PISCES. The cause of this discrepancy is the difference in models used for the minority-electron mobility in the base. For example, the SUMM mobility is 17% lower than the PISCES mobility for $N_A = 5.0 \times 10^{17} \text{ cm}^{-3}$, and 23% lower for $N_A = 1.0 \times 10^{18} \text{ cm}^{-3}$. Both simulators predict a decrease in I_C with increased base-implant dose; however the current predicted by SUMM/MMSPICE is only about 50% of that predicted by PISCES for the higher-doped device. This discrepancy is partly

attributable to the mobility models. What is more important, however, is that PISCES [Law90], unlike SUMM/MMSPICE, predicts (equivocally) a significant influence of bandgap narrowing in the QNB on I_C in the higher-doped BJT. Additional simulations show that the bandgap narrowing assumed in PISCES [Law90] tends to increase I_C in this device by about 65% over the current corresponding to no bandgap narrowing.

If we desired to match the predicted base and collector currents, we could have calibrated SUMM/MMSPICE to PISCES by tuning certain MMSPICE parameters. For example, we could reduce the difference in collector currents by increasing the value of the average electron diffusivity in the base for the MMSPICE model. We could easily match base currents by tuning MMSPICE parameters that are characterized by nonfundamental lifetimes and heavy-doping effects. By reducing the parameter JE0 (see Section 2.2), we could match the base currents at moderate bias (i.e., reduce the predicted hole injection into the emitter.) However, such parameter tuning would be counterproductive because the same uncertainties underlie the modeling of both simulators. Indeed, these results stress the inevitable need for experimentally based tuning of nonfundamental and/or ill-defined parameters in the physical models underlying all TCAD tools.

Fig. 2.4 shows β versus V_{BE} characteristics ($V_{CE} = 2$ V) predicted by SUMM/MMSPICE and PISCES for the nominal (lower-doped) device of Fig. 2.2. The characteristics have been normalized to their maximum values, which differ because of

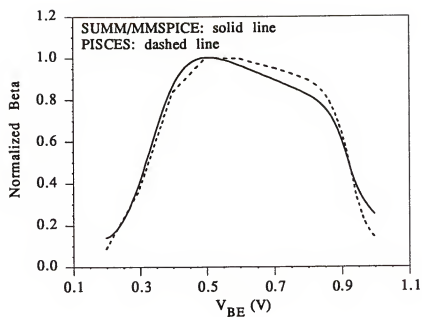


Fig. 2.4. Simulated normalized β versus V_{BE} ($V_{CE} = 2$ V) characteristics for the nominal device of Fig. 2.2.

the differences in physical modeling discussed with reference to Fig. 2.3. The predicted characteristics in Fig. 2.4 are quite similar, and serve to verify the predictive capability of the SUMM/MMSPICE charge-based BJT model. For example in high current operation, the complex effects of quasi-saturation, including base push-out, are simulated well.

Fig. 2.5 shows gain-bandwidth product GB versus I_c characteristics ($V_{CE} = 2$ V) predicted by SUMM/MMSPICE and PISCES for the nominal device of Fig. 2.2. The PISCES simulations were generated using its enhanced [Law90] small signal analysis capabilities. The MMSPICE simulations were generated using the quasi-static linearized version of the BJT model. Parameter evaluation criteria [Gre90] ensured that the variations of QNB boundaries (and hence junction space charge) with applied bias were faithfully simulated. The characteristics predicted by SUMM/MMSPICE and PISCES correspond well, and the collector currents corresponding to the peak GB values are nearly equal. This correspondence supports the characterization of the collector charge dynamics in the BJT model, which control the high- I_c GB. Discrepancies in the characteristics are due to the aforementioned differences in physical models, which influence regional charges in the BJT, including the charge in the polysilicon emitter contact. This charge, which is neglected in PISCES but physically accounted for in SUMM (see Appendix), contributes significantly to the device delay, and therefore reduces values of GB predicted by SUMM/MMSPICE.

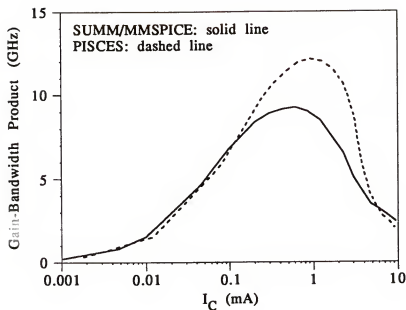


Fig. 2.5. Simulated GB versus I_C characteristics ($V_{CE} = 2$ V) for the nominal device of Fig. 2.2.

Indeed the SUMM/MMSPICE simulation of Fig. 2.5 could be more reliable than that of PISCES. Furthermore, the very high frequency ac simulation capability of PISCES is limited by convergence problems [Law90], which in fact is why we compared GB, evaluated at lower frequencies, and not the actual unity-gain frequency, f_T .

Verification and viability of the SUPREM-3/SUMM/MMSPICE system are evidenced by the simulated and measured f_T versus I_C characteristics ($V_{CE} = 3$ V) of advanced BJTs with different epitaxial-collector widths reported by Vande Voorde et al. [Van90]. They used a conventional TCAD system (Fig. 1.1(a)) comprising SUPREM-3/PISCES-II/SPICE, where PISCES-II and SPICE were linked by an empirical parameter extractor. Their simulations and measurements, of two device structures with epi thicknesses differing by $0.2\text{ }\mu\text{m}$, are shown in the inset of Fig. 2.6 (which is also Fig. 3 of [Van90].) The values of f_T plotted were normalized to about 30 GHz. The SUMM/MMSPICE system was used to simulate the same thin epi-collector structure (with the nominal doping profile in Fig. 2 of [Van90]), and a similar structure with WEPI $0.2\text{ }\mu\text{m}$ larger. Our predicted f_T versus I_C characteristics for these devices are shown in Fig. 2.6; they have been normalized to the predicted peak f_T of the nominal device, which is about 26 GHz. (Note that the nominal structure we used is identical to that used by Vande Voore et al.; however, the other structure differs slightly. In [Van90], in addition to increasing the epi thickness, the pinched base sheet resistance was

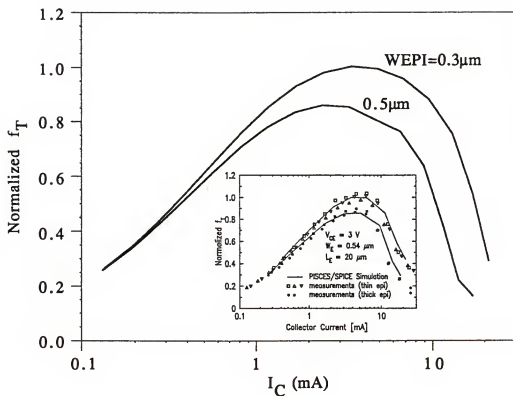


Fig. 2.6. SUMM/MMSPICE-simulated f_T versus I_C characteristics ($V_{CE} = 3$ V) for two BJTs with epi-collector widths differing by $0.2 \mu\text{m}$. The inset shows results [Van90] of corresponding measurements and PISCES/SPICE simulations with (virtually) the same devices.

increased, which we neglected. This neglect is justified because the BJTs have narrow bases ($\sim 0.065 \mu\text{m}$), and hence the change in the sheet resistance affects f_T much less than the variation in epi-collector width.) The observed reduction of f_T with increasing epi thickness is due to the increase in intrinsic collector resistance and the concomitant enhancement of quasi-saturation charge storage in the collector. Notice that the f_T versus I_C characteristics predicted by SUMM/MMSPICE agree quite well with those obtained by measurements and PISCES/SPICE simulations. Nonetheless, we stress that our results were derived without the use of optimization-based parameter extraction, and were obtained in a small fraction of the run time required for the purely numerical simulations in [Van90].

The utility of the integrated SUPREM-3/SUMM/MMSPICE system is exemplified by mixed-mode device/circuit simulations showing how ECL propagation delay is sensitive to variations in the process flow. These illustrations use the three-stage (12-transistor) ECL cascade circuit shown in Fig. 2.7. For the profiles of different base implant doses (Fig. 2.2) and the corresponding sets of model parameters evaluated by SUMM, with accountings for extrinsic collector-base and collector-substrate junction capacitances, MMSPICE simulations yielded the circuit responses shown in Fig. 2.8. Apparent in these simulations are the differences in the transient propagation delay and the output swing caused by the variation in implant dose; the delay is increased by the

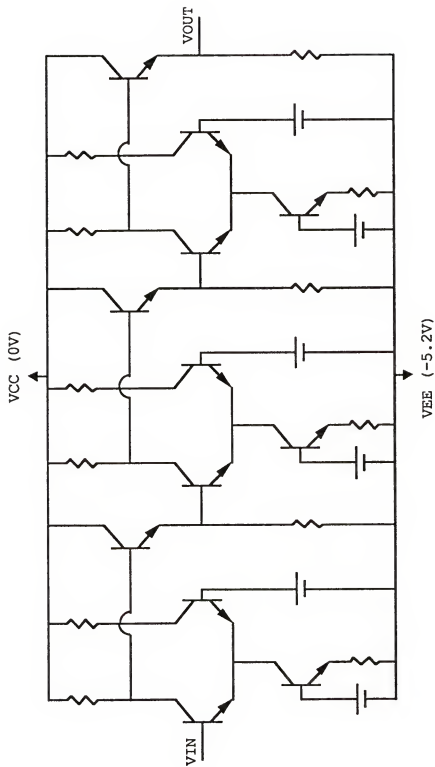


Fig. 2.7. Three-stage ECL cascade circuit for simulation.

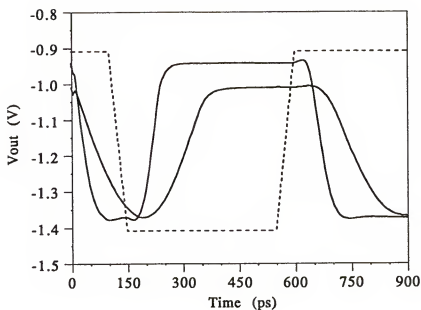


Fig. 2.8. SUMM/MMSPICE-simulated outputs of the circuit in Fig. 2.7 for the two BJT doping profiles in Fig. 2.2. The input pulse is also shown. The propagation delay is longer for the higher base-implant dose.

higher dose. These results are attributed primarily to the increases in the BJT Gummel number and base width and the concomitant changes in the current gain and base transit time. (Note that the relative changes in the ECL performances could be affected by circuit layout parasitics which were not accounted for in the simulations.)

Fig. 2.9 shows the results obtained from simulating the circuit in Fig. 2.7, for the nominal doping profile used by Vande Voorde et al. [Van90], with different values of S_p : 7×10^4 , 10^5 , and 2×10^5 cm/s, which are typical values for advanced bipolar technologies. With the model parameters corresponding to each value of S_p evaluated by SUMM, the MMSPICE simulations show variations in both the propagation delay and the output voltage swing. Note in Fig. 2.9 that the circuit delay increases with decreasing S_p . This sensitivity, which is significant in scaled technologies, is attributed primarily to the build-up of charge in the polysilicon portion of the emitter as S_p is reduced (see Appendix). We stress that if the correlations between JEO and TE, dependent on S_p as characterized in SUMM, were not properly accounted for, erroneous and misleading conclusions could be made regarding the circuit sensitivity to the emitter-contact formation, which can underlie processing instability. Furthermore, ignoring charge in the polysilicon, as is commonly done when numerical device simulators such as PISCES are used, can yield significant errors as well.

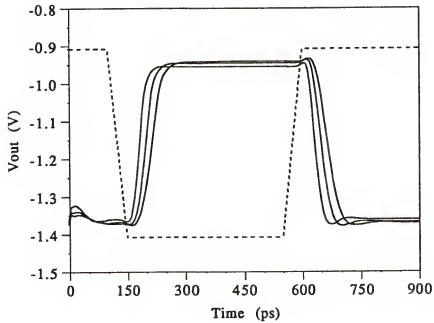


Fig. 2.9. SUMM/MMSPICE-simulated outputs of the circuit in Fig. 2.7 for different values of the surface recombination velocity at the polysilicon emitter contact: $S_p = 7 \times 10^4$, 10^5 , and 2×10^5 cm/s. The input pulse is also shown. The propagation delay increases with decreasing S_p .

2.4 Conclusions

An integrated process/device/circuit simulation system has been developed which exploits a novel alternative to conventional IC TCAD systems. This system, for advanced silicon bipolar technologies, exemplifies an approach which involves the use of numerical process simulation and pragmatic preservation of process-flow information via an application-specific, computationally efficient, seminumerical mixed-mode device/circuit simulator. The program SUMM provides the physical link between SUPREM-3 and MMSPICE by evaluating the BJT model parameters from the doping-density profile with criteria defined to ensure their physical integrity and correlation; no optimization-based (e.g., least-squares-fit) parameter extraction is used. The system was verified with both measurements and numerical device simulations using SUPREM-3/PISCES-II. The SUMM/MMSPICE system was shown to reliably predict trends reflecting process sensitivity. Discrepancies between these predictions and those made with PISCES were attributed to differences in nonfundamental and/or ill-defined physical models, and are indicative of the inevitable need to tune any TCAD tool based on measurements from a specific technology.

Simulations of ECL circuit sensitivity to process-flow variations demonstrated the utility of the SUPREM-3/SUMM/MMSPICE system in advanced bipolar TCAD. The demonstration further implies the utility of the system in manufacturing CAD (e.g., in Monte Carlo statistical

simulation for parametric yield prediction), which requires a well integrated, physics-based but computationally-efficient TCAD system.

CHAPTER 3
A MODEL OF THE SHORT-CHANNEL MOSFET
FOR PRAGMATIC MIXED-MODE DEVICE/CIRCUIT SIMULATION

3.1 Introduction

Extending the scope of the pragmatic TCAD system described in Chapter 2 to include complementary-MOSFET (CMOS) and mixed bipolar/CMOS (BiCMOS) technologies requires a predictive short-channel MOSFET circuit model that is based on structure-dependent parameters, with minimal reliance on empirical modeling. Such a model is presented in this chapter. The next chapter details the implementation of this model in SUMM/MMSPICE to create a pragmatic mixed-mode MOSFET device/circuit simulator. The development of the model for short-channel bulk-silicon MOSFETs is based in part on major revisions of an existing physical SOI MOSFET model [Vee88, Cho91], which accounts for the predominant short-channel effects. These effects include threshold voltage reduction due to charge sharing [Sze81], drain-induced channel-conductivity enhancement (DICE) [Vee88], drain-induced barrier lowering (DIBL) [Tro79], carrier velocity saturation, channel-length modulation, and generation current due to impact ionization.

Figure 3.1 shows the typical structure of the bulk enhancement mode (n-channel) MOSFET with lightly-doped drain (LDD) and source (LDS) regions. These regions are commonly

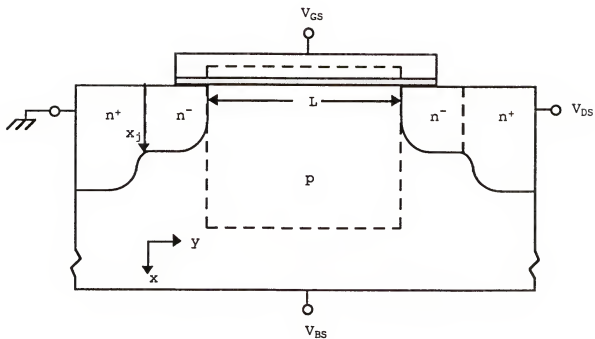


Fig. 3.1. The n-channel bulk MOSFET structure with LDD/LDS regions. The intrinsic device is demarcated by the dashed rectangle.

included in short-channel technologies to improve device reliability by reducing peak electric fields. The model developed in this chapter includes characterizations of MOSFET transport, displacement (charging), and generation currents. These characterizations are simplified by analyzing separately the intrinsic and extrinsic regions of the device. The intrinsic region is demarcated by the dashed rectangle in Fig. 3.1. The analyses of the intrinsic-region device and the LDD are linked by a Newton-like iterative solution technique. An equivalent circuit representation of the entire MOSFET is assembled in the circuit simulator SPICE2 (see Ch. 4) for device and circuit simulation.

The currents associated with the intrinsic device are derived from characterizations of the intrinsic-region charges. These derivations are simplified by considering two distinct conditions for which either drift or diffusion channel transport current predominates, viz., weak or strong surface inversion. Due to the complexity of modeling moderate inversion, which is characterized by significant drift and diffusion transport currents along the entire channel, interpolation schemes are used for computing moderate-inversion currents. To compute moderate-inversion transport current a cubic spline is used to interpolate between weak-inversion (diffusion) and strong-inversion (drift) channel currents. The spline ensures the continuities of channel current (I_{CH}) and transconductance (dI_{CH}/dV_{GS}). We found that linear interpolation was adequate for computing moderate-

inversion displacement current and the multiplication factor used to calculate moderate-inversion generation current. The former is found from interpolation between weak- and strong-inversion regional charges, while the latter is computed via interpolation of physical components that determine the peak channel electric field.

3.2 Determining the Region of Operation

In this section we derive the gate-to-source threshold voltages that define the ranges of applied biases for which the MOSFET operates with weak-, moderate-, or strong-surface inversion. In conventional MOSFET circuit models that rely on cubic-spline interpolation to evaluate moderate-inversion current, e.g., [She87], the boundaries between weak and strong inversion are used as fitting parameters. In our model these boundaries are uniquely defined by the device structure.

The lower-limit surface potential in strong inversion, Ψ_{ss} , is defined using Tsividis's analysis [Tsi82] of the one-dimensional (1-D) Poisson equation in the MOSFET body under the gate,

$$\frac{\partial^2 \Psi}{\partial x^2} = - \frac{\rho(x)}{\epsilon_s} \quad (3.1)$$

where ϵ_s is the silicon permittivity and $\rho(x)$ is the total space-charge density, with assumed Boltzmann distribution

[Sze81] of electrons and holes. For a uniformly doped (N_B) p-type semiconductor, the total space charge, the ionized impurity and majority (hole) charge, and the inversion-layer charge per unit area (Q_s , Q_b , Q_c , respectively) are found from (3.1) to be

$$Q_s = -Q_r \sqrt{\beta \psi_s - 1 + e^{\beta(\psi_s - \psi_i)}}, \quad (3.2)$$

$$Q_b = -Q_r \sqrt{\beta \psi_s - 1}, \quad (3.3)$$

and

$$Q_c = Q_s - Q_b, \quad (3.4)$$

where ψ_s is the silicon surface potential, $\psi_i = 2V_{th} \ln(N_B/n_i) - V_{BS}$, and $\beta = V_{th}^{-1}$. V_{th} is the thermal voltage (kT/q). Q_r is given by

$$Q_r = \sqrt{2q\epsilon_s N_B \beta^{-1}}. \quad (3.5)$$

The corresponding capacitances, $-dQ/d\psi_s$, are

$$C_s = C_r \frac{1 + e^{\beta(\psi_s - \psi_i)}}{2\sqrt{\beta \psi_s - 1 + e^{\beta(\psi_s - \psi_i)}}}, \quad (3.6)$$

$$C_b = \frac{C_r}{2\sqrt{\beta \psi_s - 1}}, \quad (3.7)$$

and

$$C_c = C_s - C_b. \quad (3.8)$$

where $C_r = \beta Q_r$.

The surface potential ψ_{ss} is evaluated from an analysis of the (two-terminal) MOS structure [Tsi82]. The summation of potentials across the device along with Gauss's law yields

$$V_{GB} - V_{FB} = \psi_s - \frac{Q_c + Q_b}{C_{ox}}, \quad (3.9)$$

where V_{FB} is the flatband voltage. Equation (3.9), with (3.2)-(3.5), indicate that deep in strong inversion ψ_s and Q_b are approximately constant [Sze81]. This then suggests that the degree to which dQ_c/dV_{GB} is approximately $-C_{ox}$ ($= -\epsilon_{ox}/t_{ox}$) is a physical criterion for determining ψ_{ss} . In general, dQ_c/dV_{GB} ($= dQ_c/d\psi_s * d\psi_s/dV_{GB} = -C_c * d\psi_s/dV_{GB}$) is expressed as

$$\frac{dQ_c}{dV_{GB}} = - \frac{C_{ox}C_c}{C_{ox}+C_b+C_c}, \quad (3.10)$$

using (3.9). Based on (3.10) ψ_{ss} is defined as the value of ψ_s for which

$$C_c = 10(C_b + C_{ox}), \quad (3.11)$$

and therefore dQ_c/dV_{GB} is $-C_{ox}/1.1$ ($\equiv -C_{ox}$). Now (3.6)-(3.8) in (3.11) yield

$$\frac{1 + e^{\beta(\psi_{ss} - \psi_i)}}{2\sqrt{\beta\psi_{ss} - 1 + e^{\beta(\psi_{ss} - \psi_i)}}} - \frac{11}{2\sqrt{\beta\psi_{ss} - 1}} - 10\frac{C_{ox}}{C_r} = 0, \quad (3.12)$$

which is used to compute ψ_{ss} via Newton-Raphson iteration [Pre89]. The value of V_{GS} ($= V_{GB} + V_{BS}$) corresponding to ψ_{ss} (for $V_{DS} = 0$), V_{TS0} , is found from (3.9) with (3.2):

$$V_{TS0} = V_{BS} + V_{FB} + \psi_{ss} - \frac{Q_r}{C_{ox}} \sqrt{\beta\psi_{ss} - 1 + e^{\beta(\psi_{ss} - \psi_i)}}. \quad (3.13)$$

Note ψ_{ss} is greater than ψ_i , which is commonly assumed to indicate strong inversion.

The transition between weak and moderate inversion is analogously derived to be $\psi_{sW} \equiv \psi_i$ [Tsi82]. However, we found that due to accuracy limitations of the cubic spline, a value of ψ_{sW} less than ψ_i was needed; simulations with the completed model suggested the value $\psi_{sW} = \psi_i - 5V_{th}$, which we will use. The value of V_{GS} corresponding to this smaller ψ_{sW} (for $V_{DS} = 0$), V_{TW0} , is found from (3.9) with Q_c negligible:

$$V_{TW0} = V_{BS} + V_{FB} + \psi_i - 5V_{th} - \frac{Q_b}{C_{ox}} \quad (3.14)$$

where Q_b is given by (3.3) with $\psi_s = \psi_{sW}$.

3.2.1 Charge Sharing

For short-channel MOSFETs the electric fields due to the source and drain junctions ($V_{DS} = 0$) affect gate control of the device, requiring us to consider modifications of the threshold voltages in (3.13) and (3.14) which were derived from the 1-D Poisson equation (3.1). This two-dimensional (2-D) electric field effect, which can cause the MOSFET to produce more channel current than is predicted by long-channel models [Sze81], differs fundamentally in weak and strong inversion. Shown in Fig. 3.2 are the electric field vectors in a portion of a PISCES [Tec89]-simulated n-MOSFET structure for both strong inversion (Fig. 3.2(a)) and weak inversion (Fig. 3.2(b)). For these simulations $V_{DS} = 0$, and the technology parameters used were $L = 0.5 \mu\text{m}$, $N_B = 1.5 \times 10^{17} \text{ cm}^{-3}$, $t_{ox} = 13.3 \text{ nm}$ and $x_j = 0.3 \mu\text{m}$. The bounded region along the top of the structure represents the gate oxide, and the dashed lines indicate the source and drain junctions. The encroachment of junction electric field under the gate oxide produces the short-channel effect referred to as charge sharing.

As illustrated in Fig. 3.2(a) the oxide field under the gate in strong inversion is fairly uniform, indicating a nearly constant surface potential. We therefore account for strong-inversion charge sharing using a conventional modification of (3.13) based on the charge conservation principle introduced by Yau [Yau74]. Yau's simple model relies on a geometric analysis of the 2-D electric field to

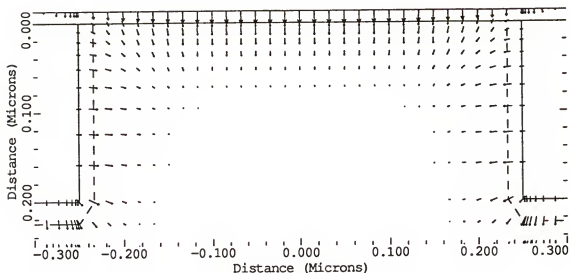
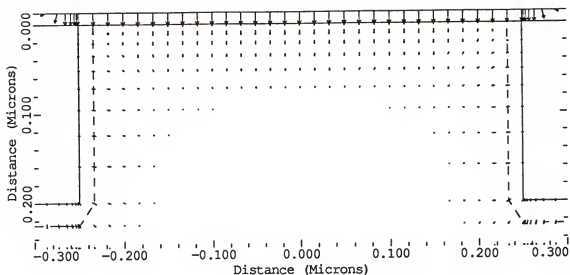


Fig. 3.2. A portion of a PISCES-simulated n-MOSFET structure with the electric field vectors for (a) strong inversion and (b) weak inversion. For each case the lengths of the vectors have been normalized by the highest field value. The bounded region along the top of the structure represents the gate oxide. Notice in (b) the oxide field is severely nonuniform relative to that in (a) near the source and drain junctions (dashed lines).

define an effective gate-controlled (areal) body charge density

$$Q_{b(\text{eff})} = Q_b \left[1 - \frac{x_j}{L} \left(\sqrt{1 + \frac{2x_{d0}}{x_j}} - 1 \right) \right] \Delta Q_b f \quad (3.15)$$

where L and x_j are the channel length and drain/source junction depth, respectively; and x_{d0} is the depletion-region depth at the center of the channel corresponding to V_{GS} and V_{BS} with $V_{DS} = 0$. This depth is defined (for both weak and strong inversion) by equating Q_b in (3.3) to $-qN_B x_{d0}$:

$$x_{d0} = \sqrt{\frac{2\epsilon_{s1}(\Psi_s - V_{th})}{qN_B}} \quad , \quad (3.16)$$

and is evaluated in strong inversion using $\Psi_s = \Psi_{sS}$. Note that (3.15) gives a representative $|Q_{b(\text{eff})}| < |Q_b|$ irrespective of the relative values of x_j and x_{d0} . The corresponding reduction of V_{TS0} is accounted for in our model through the reduction of Q_r (given by (3.5)) in (3.13):

$$Q_{r(\text{eff})} = Q_r f \quad (3.17)$$

where the factor f (< 1) is defined in (3.15). (Note in accordance with Yau's model we do not include f in (3.12) for the evaluation of Ψ_{sS} .)

It is obvious in Fig. 3.2(b) that charge sharing in weak inversion is complicated by the pronounced nonuniformity in

the oxide field under the gate near the source and drain. (This effect is separate from field fringing.) Consequently, the "long-channel" assumption of a constant surface potential is grossly inadequate. Indeed the influence of the junctions on weak-inversion current is more accurately described as channel-length modulation, even for $V_{DS} = 0$, rather than charge sharing. Notice though in Fig. 3.2(b) that there is a region near the gate center where the influence of the drain and source is negligible, which is typical for scaled devices. Therefore we use, unmodified, (3.14) to express V_{TWO} , and account for the nonconstant surface potential in the weak-inversion channel-length modulation model derived later.

3.2.2 DIBL and DICE

For short-channel MOSFETs, applied drain-source voltages enhance the channel current through the mechanisms of DIBL [Tro78] in weak inversion and DICE [Vee88] in strong inversion. In the former, V_{DS} modulates the surface potential directly, while DICE is characterized by indirect modulation of channel charge as dictated by the 2-D Poisson equation under the gate. We need to account for the impact of these two mechanisms on the threshold voltages V_{TWO} and V_{TSO} (given by (3.13) and (3.14)).

To properly model 2-D electric field effects such DIBL and DICE it is necessary to analyze the 2-D Poisson equation. To simplify this analysis we perform separate yet coupled low longitudinal-field and high longitudinal-field analyses. The

high-field analysis employs Gauss's law in two dimensions to form the basis of the channel-length modulation model, and is addressed later. The low-field analysis (no "pinch-off") forms the basis of the DIBL/DICE model. For that analysis of the MOSFET illustrated in Fig. 3.3, we focus on the area under the gate defined by $0 \leq y \leq L$ and $0 \leq x \leq x_{d0}$, where $x = 0$ lies just below the conduction channel; x_{d0} is the depletion region depth (at the channel center) defined in (3.16) corresponding V_{GS} and V_{BS} with $V_{DS} = 0$. In strong inversion the surface potential Ψ_{s0} used to evaluate (3.16) is the V_{GS} -independent Ψ_{ss} ($\gg V_{th}$), while in weak inversion Ψ_{s0} is computed from (3.9) with $|Q_{c0}| \ll |Q_b|$,

$$\Psi_{s0} = V_{GS} - V_{BS} - V_{FB} + \frac{Q_b}{C_{ox}} , \quad (3.18)$$

with Q_b expressed in (3.3). In the region $0 \leq y \leq L$ and $0 \leq x \leq x_{d0}$ the 2-D Poisson equation simplifies to Laplace's equation for the incremental change in potential due to V_{DS} :

$$\frac{\partial^2 \Delta\psi}{\partial x^2} + \frac{\partial^2 \Delta\psi}{\partial y^2} = 0 . \quad (3.19)$$

Numerical simulations of short-channel devices suggest that for low (but not negligible) longitudinal-fields we can extrapolate from the gradual-channel approximation (GCA) [Sze81], and solve (3.19) by a quasi-separation of variables:

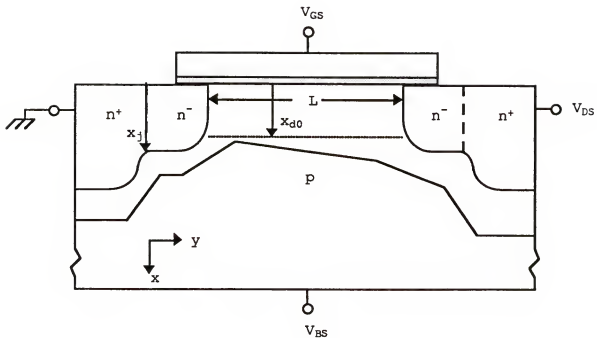


Fig. 3.3. The n-channel MOSFET structure with a representative depletion-layer boundary that reflects the prevalent 2-D electric field in the short-channel device. The depletion depth x_{d0} is defined under the center of the gate for $V_{DS} = 0$.

$$\frac{\partial^2 \Delta\psi}{\partial x^2} \equiv -\frac{\partial^2 \Delta\psi}{\partial y^2} \equiv -\eta \quad (3.20)$$

where $\eta = (2/L^2)V_{DS}$ [Vee88]. This first-order extrapolation reduces to the GCA for large L .

Integrating (3.20) over x yields

$$\Delta E_s(y) = \frac{\Delta\psi_s(y) - \Delta\psi(x_{d0}, y)}{x_{d0}} - \frac{\eta x_{d0}}{2}, \quad (3.21)$$

which expresses the change in surface transverse electric field as a function of the surface potential change, the change in potential $\Delta\psi(x_{d0}, y)$, and V_{DS} . Based on physical insight (3.21) can be simplified. In weak inversion we use the approximation $\Delta\psi_s(y) \equiv \Delta\psi(x_{d0}, y)$, which is consistent with the case of weak gate control of the charge in the silicon near the body-drain junction. In strong inversion however we neglect $\Delta\psi(x_{d0}, y)$ relative to $\Delta\psi_s(y)$; this approximation is equivalent to assuming a constant depletion capacitance along the channel.

Assuming vertical fields in the oxide, we apply Gauss's law at the Si-SiO₂ interface with (3.21) and the mentioned strong-inversion approximation to get

$$\Delta Q_c(y) = \left(C_{ox} + \frac{\epsilon_s}{x_{d0}} \right) \Delta\psi_s(y) - \frac{\epsilon_s \eta x_{d0}}{2}, \quad (3.22)$$

which expresses the subsequent change in channel charge in strong inversion. The first term in (3.22) represents the conventional long-channel result with $(\epsilon_s/x_{d0})\Delta\psi_s(y)$ providing an approximation for the body effect. The second term in (3.22) expresses the channel charge enhancement due to the drain-source bias, which is the effect called DICE.

In weak inversion Gauss's law with the mentioned approximation in (3.21) yields

$$\Delta Q_c(y) = C_{ox}\Delta\psi_s(y) - \frac{\epsilon_s\eta x_{d0}}{2} . \quad (3.23)$$

It can be shown that $\Delta Q_c(y)$ is much less than the terms on the right-hand side of (3.23), so the effect of DIBL is seen through its influence on the surface potential:

$$\Delta\psi_s^{DIBL} \equiv \frac{\epsilon_s\eta x_{d0}}{2C_{ox}} , \quad (3.24)$$

which is independent of y .

In weak inversion the influence of DIBL on V_{TW0} is due to the direct dependence of ψ_s on V_{DS} through (3.24), producing a V_{DS} -dependent threshold voltage V_{TW} , which is less than V_{TW0} :

$$V_{TW} \equiv V_{TW0} - \Delta\psi_s^{DIBL} = V_{TW0} - \Delta V_{TW} \quad (3.25)$$

where $\psi_{s0} = \psi_{sW} = \psi_i - 5V_{th}$ is used to evaluate ΔV_{TW} .

Via DICE the strong-inversion threshold will also be reduced. An expression for this reduction is easily derived from an analysis of the charge at the source end of the channel ($y = 0$) where $\Delta\psi_s(y) = 0$. The DICE model of (3.22), with (3.9) expressed at $V_{GB} \equiv V_{GS} - V_{BS} = V_{TS} - V_{BS}$, produces a V_{DS} -dependent threshold voltage

$$V_{TS} = V_{TS0} - \frac{\epsilon_s \eta x_{d0}}{2C_{ox}} \equiv V_{TS0} - \Delta V_{TS} \quad (3.26)$$

where $\psi_{s0} = \psi_{ss}$ is used to evaluate ΔV_{TS} . (Note that ΔV_{TW} and ΔV_{TS} are functions the voltage drops across the LDD and LDS due to their dependencies on V_{BS} and V_{DS} . These voltage drops will be discussed later.)

With regard to measured device characteristics, the conventional extrapolated threshold voltage is greater than V_{TW} , but less than V_{TS} . It can be shown that this threshold voltage experiences the same shift with V_{DS} as V_{TS} .

3.3 Strong-Inversion Channel Transport Current

For strong surface inversion, the channel transport current is predominantly due to carrier drift, and can be expressed

$$I_{CH} = -WvQ_c \quad (3.27)$$

where v is electron drift velocity and Q_c is the channel charge (areal) density. The method for evaluating this charge density is to compute separately its value for $V_{DS} = 0$, Q_{c0} , and the perturbation due to $V_{DS} > 0$, $\Delta Q_c(y)$. The latter is computed from (3.22). Q_{c0} is expressed from (3.9) with Q_b replaced by $Q_{b(eff)}$ in (3.15):

$$Q_{c0} = -C_{ox}(V_{GS} - V_{BS} - V_{FB} - \psi_{ss} + \frac{Q_{b(eff)}}{C_{ox}}) , \quad (3.28)$$

where the biases are now referenced to the source, and ψ_{s0} is assumed constant and equal to ψ_{ss} (from (3.12)). Note that this value of ψ_{s0} is greater than the conventionally used value of ψ_i as a result of the physical criterion for defining strong inversion in (3.11).

In the short-channel MOSFET high transverse and longitudinal electric fields in the channel can cause significant nonlinearity in the carrier velocity-field characteristic. For sufficiently large longitudinal field $|E_y| = d\psi_s/dy$, the carrier velocity will saturate at v_{sat} ($\cong 10^7$ cm/s for n-MOSFETs), manifesting channel-length modulation. The carrier velocity, v , in (3.27) is modeled by [Vee88]

$$v(y) = \frac{\mu_{eff}|E_y|}{1 + \frac{\mu_{eff}|E_y|}{2v_{sat}}} , \quad v(y) \leq v_{sat} \quad (3.29)$$

$$v(y) = v_{sat} , \quad \text{otherwise.}$$

3.3.1 Triode Current

In (3.29), μ_{eff} is the low-(longitudinal-)field mobility, and is dependent on the transverse field, E_x , in the channel. This dependence is modeled [Vee88] in terms of the average $\bar{E}_x(y)$ along the channel:

$$\mu_{\text{eff}} = \frac{\mu_{n0}}{1 + \Theta \bar{E}_x(y)}, \quad (3.30)$$

where Θ is an empirical constant.

The average transverse field is defined as $\bar{E}_x(y) \triangleq \bar{E}_{x0} + \overline{\Delta E_x}(y)$, where \bar{E}_{x0} is the average field due to V_{GS} and V_{BS} with $V_{DS} = 0$, and $\overline{\Delta E_x}(y)$ represents the change in this field due to $V_{DS} > 0$. \bar{E}_{x0} is expressed as [Vee88]

$$\bar{E}_{x0} = -\frac{1}{\epsilon_s} \left(\frac{Q_{c0}}{2} + Q_{b(\text{eff})} \right), \quad (3.31)$$

which is evaluated using (3.15) and (3.28). The field $\overline{\Delta E_x}(y)$ is similarly evaluated,

$$\overline{\Delta E_x}(y) = -\frac{1}{\epsilon_s} \left(\frac{\Delta Q_c(y)}{2} + \Delta Q_{b(\text{eff})}(y) \right). \quad (3.32)$$

The first term in (3.32) can be expressed using (3.22) from the DICE analysis. For the second term, (3.21) with Gauss's law and the neglect of $\Delta\psi(x_{d0}, y)$ gives the change in gate-controlled body depletion charge:

$$\Delta Q_{b(eff)}(y) = -\frac{\epsilon_s}{x_{d0}} \Delta \Psi_s(y) + \frac{\epsilon_s x_{d0} \eta}{2}. \quad (3.33)$$

The first term in (3.33) represents the increase in charge, approximating the long-channel solution using a constant depletion capacitance; this is a fair approximation in the low-field region where the velocity is not saturated. The second term in (3.33) expresses the reduction in depletion charge that results from the increased inversion charge caused by the drain bias.

Using (3.29) with (3.27), we obtain the steady-state channel current for $v(L) < v_{sat}$ via integration over the length of the channel:

$$I_{CH} = \frac{\bar{\mu}_{eff} [Q_c^2(0) - Q_c^2(L)]}{2L(C_{ox} + \frac{\epsilon_s}{x_{d0}}) (1 + \frac{\bar{\mu}_{eff} V_{DS}}{2v_{sat}L})} \quad (3.34)$$

where

$$\bar{\mu}_{eff} = \frac{\mu}{1 - BV_{DS}/2} \quad (3.35)$$

with

$$\mu = \frac{\mu_{n0}}{1 - \Theta \left(\frac{Q_{c0}}{2\epsilon_s} + \frac{Q_{b(eff)}}{\epsilon_s} + \frac{\eta x_{d0}}{4} \right)} \quad (3.36)$$

and

$$B = \frac{\Theta \frac{C_{ox} - \frac{\epsilon_s}{x_{d0}}}{2\epsilon_s}}{1 - \Theta \left(\frac{Q_{c0}}{2\epsilon_s} + \frac{Q_{b(eff)}}{\epsilon_s} + \frac{\eta x_{d0}}{4} \right)} \quad (3.37)$$

Note to obtain (3.35) the approximation $\int_0^L \Delta\psi_s dy \cong V_{DS}L/2$ was used in accordance with a representative linear dependence of $\Delta\psi_s$ on y ; this approximation is not assumed anywhere else in the model, and in no way implies that a linear dependence dictates (3.34).

3.3.2 Saturation Current

In strong-inversion saturation the channel is divided into a region where the carrier velocity is saturated (drain-side) and a region where the velocity is field-dependent (source-side). The source-side and drain-side regions are separated at the modulated channel length, $L_e = L - \Delta L$. The saturated channel current is derived like (3.34) but here the integration is over the modulated channel length ($0 \leq y \leq L_e$) for which $\Delta\psi_s(y)$ varies from 0 to $V_{DS(eff)}$:

$$I_{CH(sat)} = -WQ_C(L_e) v_{sat} = \frac{W\bar{\mu}_{eff}' [Q_C^2(0) - Q_C^2(L_e)]}{2L(C_{ox} + \frac{\epsilon_s}{x_{d0}}) (1 + \frac{\bar{\mu}_{eff}' V_{DS(eff)}}{2v_{sat}L_e})} \quad (3.38)$$

where

$$\bar{\mu}_{eff}' = \frac{\mu}{1 - BV_{DS(eff)}/2} . \quad (3.39)$$

The saturated-current analysis is completed with an evaluation of $V_{DS(eff)}$ and L_e . For this we prefer not to use the solution to the 2-D Poisson equation derived in Section 3.2.2, because it is based on a low-longitudinal electric field approximation. To evaluate $V_{DS(eff)}$ and L_e we use an extended version of the channel-length modulation analysis presented by El-Mansy [Man77]. That analysis is based on an approximate solution to the 2-D Poisson equation in a high-field region.

To characterize L_e we again focus on the MOSFET structure of Fig. 3.3, and apply Gauss's law to an incremental portion (dy) of the region defined by $L_e \leq y \leq L$ and $0 \leq x \leq x_{d0}$, where $x = 0$ lies just above the Si-SiO₂ interface. The Gaussian surface we have selected is bounded below by x_{d0} , the depletion-region depth evaluated from (3.16) using $\psi_{s0} = \psi_{ss}$ ($V_{DS} = 0$). Analogous to the treatment in [Man77], the application of Gauss's law yields

$$\begin{aligned} \Delta Q_c(y) dy = & \epsilon_s \Delta E_x(x_{d0}, y) dy - \epsilon_{ox} \Delta E_{ox}(y) dy \\ & + \epsilon_s \int_0^{x_{d0}} \Delta E_y(x, y+dy) dx - \epsilon_s \int_0^{x_{d0}} \Delta E_y(x, y) dx \end{aligned} \quad (3.40)$$

where the Δ 's represent changes in variables resulting from the application of V_{DS} . Note within the assumed Gaussian surface defined by x_{d0} that the only charge which can change with V_{DS} is Q_c , the channel electron charge (density).

Because both the current and carrier velocity are constant throughout the high-field drain-side region, the channel charge density must also be constant. Therefore, the left-hand side of (3.40) can be expressed as

$$\Delta Q_c(y) dy = \Delta Q_c(L_e) dy = (C_{ox} + \frac{\epsilon_s}{x_{d0}}) \Delta \psi_s(L_e) dy - \frac{\epsilon_s x_{d0} \eta}{2} dy \quad (3.41)$$

with the results of the DICE analysis. Note $\Delta \psi_s(L_e) = V_{DS(eff)}$, and η remains a function of V_{DS} and L .

The DICE analysis in Section 3.2.2 also provides an evaluation of the transverse field at x_{d0} in the first term on the right-hand side of (3.40). Integrating (3.20) over x , and using (3.21) with the neglect of $\Delta \psi(x_{d0}, y)$ yields

$$\Delta E_x(x_{d0}, y) \equiv \frac{\Delta \psi_s(y)}{x_{d0}} + \frac{\eta x_{d0}}{2} . \quad (3.42)$$

Note that (3.42) was derived from the low-field analysis, though we use it as an approximation for this field in (3.40).

The second term on the right-hand side of (3.40), the change in the oxide field displacement, can be expressed

$$\epsilon_{ox}\Delta E_{ox}(y)dy = \epsilon_{ox}\frac{\Delta\psi_{ox}(y)}{t_{ox}}dy = -C_{ox}\Delta\psi_s(y)dy \quad (3.43)$$

For the third and fourth terms of (3.40), we make the simplifying assumption that $\Delta E_y(y) = -d\Delta\psi/dy$ is independent of x and thus express these terms as a function of the incremental potential at the Si-SiO₂ interface, $\Delta\psi_s(y)$:

$$\begin{aligned} \epsilon_s \int_0^{x_{d0}} \Delta E_y(x, y+dy) dx - \epsilon_s \int_0^{x_{d0}} \Delta E_y(x, y) dx \\ \cong \epsilon_s x_{d0} d\Delta E_y(y) = -\epsilon_s x_{d0} \frac{d^2 \Delta\psi_s}{dy^2} dy \end{aligned} \quad (3.44)$$

With equations (3.41)-(3.44), (3.40) can be rewritten in terms of incremental potential change due to V_{DS} :

$$\frac{d^2 \Delta\psi_s}{dy^2} = \frac{C_{ox} + \epsilon_s/x_{d0}}{\epsilon_s x_{d0}} [\Delta\psi_s - \Delta\psi_s(L_e)] + \eta \quad (3.45)$$

The evaluation of the solution of (3.45) at $y = L$ with $\Delta\psi_s(L) = V_{DS}$ and $-\Delta E_y(L_e) = -E_{sat} = 2v_{sat}/\mu_{eff}$, obtained from (3.29), yields

$$V_{DS} - V_{DS(eff)} = \frac{2v_{sat}l_c}{\mu_{eff}} \sinh\left(\frac{L - L_e}{l_c}\right) + \eta l_c^2 \left[\cosh\left(\frac{L - L_e}{l_c}\right) - 1 \right] \quad (3.46)$$

where

$$l_c = \sqrt{\frac{\epsilon_s x_{d0}}{C_{ox} + \epsilon_s / x_{d0}}} . \quad (3.47)$$

Unlike the result in [Man77] (3.46) accounts for the 2-D electric field distribution in the source-side region via the inclusion of the DICE effect in the second term on the right-hand side. For properly scaled MOSFETs DICE is suppressed, and that term affects (3.46) minimally. Equations (3.38) and (3.46) provide an implicit solution for L_e and $V_{DS(eff)}$, and hence $I_{CH(sat)}$.

3.4 Weak-Inversion Channel Transport Current

For weak surface inversion there exists a region under the gate where the channel current is predominantly diffusion and can be expressed as

$$I_{CH} = W V_{th} \mu_n \frac{-\partial Q_c(y)}{\partial y} . \quad (3.48)$$

To compute the electron charge density $Q_c(y)$ in (3.48) the Boltzmann distribution [Sze81] of electrons is integrated over the channel thickness ($0 \leq x \leq x_1(y)$). In weak inversion the electrostatic potential across the channel thickness varies only a few (m) V_{th} , so $Q_c(y)$ can be derived as [Cho91]

$$-Q_c(y) \cong \frac{q n_i^2}{N_B} \frac{1}{E_x} \int_{\psi_s - mV_{th}}^{\psi_s} \exp\left(\frac{\psi - V(y)}{V_{th}}\right) d\psi$$

$$\equiv \frac{q n_i^2}{N_B} \frac{V_{th}}{E_x} \exp\left(\frac{\psi_s - V(y)}{V_{th}}\right) \quad (3.49)$$

where \bar{E}_x is an average transverse field in the channel, and $V(y)$ is the separation of electron and hole quasi-Fermi levels (QFL's) at position y along the channel. Note that it is not possible to compute the derivative of (3.49) to find the current in (3.48). I_{CH} can be found by integrating (3.48) over a length of the channel ($y_s \leq y \leq L - y_d$) bound by locations where $Q_c(y)$ is well defined:

$$I_{CH} = \frac{W}{L_e} \bar{\mu}_n V_{th} [Q_c(L - y_d) - Q_c(y_s)], \quad (3.50)$$

where y_s is near the source and y_d is near the drain. Defining y_s and y_d and evaluating (3.50) are complicated by the influence of the source and drain junctions on the silicon surface potential, which is more pronounced in weak inversion than in strong inversion. This influence is reflected by the nonuniform electric field in the oxide under the gate for the MOSFET illustrated in Fig. 3.2(b). The effect on potential on both sides of the channel, which is even significant for $V_{DS} = 0$, manifests as channel-length modulation. We therefore define the length $L_e = L - y_s - y_d = L - \Delta L$ as the modulated channel length, which differs fundamentally from L_e derived for strong inversion.

In conventional compact circuit models for short-channel MOSFETs (e.g., [Cha87] and [She87]), channel-length

modulation in weak inversion is either modeled simply by a solution to Poisson's equation in one (longitudinal) dimension to express y_s and y_d [Tay78], or not modeled explicitly, but rather accounted for in an empirical V_{DS} dependence [She87] of the subthreshold channel current, which does not separate the effect from DIBL. The former approach, which neglects the transverse electric field, is non-physical and models based on it, like those based on the latter approach, tend to be strongly dependent on empirical fitting parameters. As a result of the empiricism most subthreshold models are not physically insightful, and their strong dependence on fitting parameters complicates parameter extraction and sacrifices reliable MOSFET circuit simulation.

This new model contains a quasi-2-D analysis of channel-length modulation in short-channel MOSFETs which inherently accounts for DIBL, and which reveals that in weak inversion the modulated channel length can be approximated as independent of V_{DS} , thereby simplifying (and correcting) MOSFET circuit simulation. Numerical simulations are used for verification. The physical model further provides insight regarding how channel-length modulation (in weak inversion) scales with structural parameters of the device.

Our analysis modifies the quasi-2-D treatment by El-Mansy and Boothroyd [Man77] of current saturation in strong inversion to improve the widely used subthreshold channel-length modulation model of Taylor [Tay78]. In the analysis of [Man77], the unmodulated channel length (L) is sectioned into

two regions which are treated separately: a 2-D analysis is used in the high-field drain-side region and a conventional MOSFET analysis based on the gradual-channel approximation (GCA) is applied in the source-side region. While the GCA is applicable to the long-channel device, it is invalid for short-channel (submicron) devices due to the 2-D electric field under the gate, as evidenced for example by the manifestation of DIBL in weak inversion [Tro79]. The conventional drain-side 2-D analysis [Man77], which assumes strong inversion, cannot be directly applied to weak inversion, requiring enhancements to physically model the effects of DIBL as well as to account for the prevalent diffusion current in the channel.

As in [Tay78] we define the modulated channel length in weak inversion as the portion of the channel where carrier diffusion is predominant and where the longitudinal electric field is small relative to the fields closer to the source and drain junctions. Using an enhancement of the analysis of [Man77], we derive a revision of the 1-D estimations of y_s and y_d in [Tay78] that accounts for the 2-D electric field under the gate.

We focus on the MOSFET structure illustrated in Fig. 3.3 to characterize y_d , and apply Gauss's law to an incremental portion (dy) of the region defined by $L - y_d \leq y \leq L$ and $0 \leq x \leq x_{d0}$, where $x = 0$ lies just above the Si-SiO₂ interface. The Gaussian surface we have selected is bounded below by x_{d0} , the depletion-region depth under the center of the gate

corresponding to V_{GS} and V_{BS} when $V_{DS} = 0$, even though the depletion in the region of interest extends below this depth. With the surface potential near the center of the gate defined by V_{GS} and V_{BS} , x_{d0} is evaluated from (3.16). Analogous to the treatment in [Man77], the application of Gauss's law yields

$$\epsilon_s x_{d0} \frac{d^2 \Delta\psi_s(y)}{dy^2} - C_{ox} \Delta\psi_s(y) \cong -\Delta Q_c(y) + \epsilon_s \Delta E_x(x_{d0}, y). \quad (3.51)$$

where the Δ 's represent changes in variables resulting from the application of V_{DS} . In (3.51) we have made the simplifying assumption that $\Delta E_y(y) = -d\Delta\psi/dy$ is independent of x and thus expressed the first term as a function of the incremental potential at the Si-SiO₂ interface, $\Delta\psi_s(y)$. This potential in the second term characterizes the change in the normal electric displacement, $\epsilon_{ox} \Delta E_{ox}(y)$, along the $x = 0$ boundary. Note within the assumed Gaussian surface defined by x_{d0} that the only charge which can change with V_{DS} is Q_c , the channel electron charge (density). This change is represented by the third term in (3.51). In strong inversion this term overwhelms the last term in (3.51), the displacement normal to the bottom of the Gaussian surface, but in weak inversion this displacement is predominant. We use the field $\Delta E_x(x_{d0}, y)$ derived from the DIBL analysis (Section 3.2.2) of the low longitudinal-field region ($y_s \leq y \leq L - y_d$) as an approximation for that field in this high-field region.

Integrating (3.20) over x and using (3.21) with the approximation $\Delta\psi_s(y) \equiv \Delta\psi(x_{d0}, y)$ yields the needed field:

$$\Delta E_x(x_{d0}, y) \equiv \frac{\eta x_{d0}}{2} . \quad (3.52)$$

Therefore (3.51) can be rewritten as

$$\frac{d^2 \Delta\psi_s}{dy^2} \equiv \frac{C_{ox}}{\epsilon_s x_{d0}} \Delta\psi_s + \frac{\eta}{2} . \quad (3.53)$$

To solve (3.53) along the length of channel $L - y_d \leq y \leq L$, the boundary conditions $\Delta E_y(L - y_d)$ and $\Delta\psi_s(L - y_d)$ must be evaluated. The latter is expressed by (3.24) of the DIBL analysis. The evaluation of $\Delta E_y(L - y_d)$ however requires new insight. To be consistent with the assumption of predominant diffusion in (3.50), this incremental field must be relatively small; i.e., at $y = L - y_d$, $d\phi_{Fn}/dy \gg -E_y$ where ϕ_{Fn} is the electron quasi-Fermi potential. We approximate this condition at $(L - y_d)$ with $\Delta E_y(L - y_d) \equiv 0$.

Evaluating the resulting solution of (3.53) at $y = L$ gives

$$V_{DS} = \eta l_c^2 \left[\cosh\left(\frac{y_d}{l_c}\right) - \frac{1}{2} \right] \quad (3.54)$$

where $l_c = \sqrt{\epsilon_s x_{d0} / C_{ox}}$. From (3.54) we can solve for y_d :

$$y_d \equiv l_c \ln \left(\frac{L^2}{l_c^2} \right). \quad (3.55)$$

This surprising result indicates that in weak inversion the drain-side channel length modulation is a structurally-dependent $y_d(V_{GS}, V_{BS})$ which does not vary with V_{DS} . (For structures characterized by $x_j < x_{d0}$ the dependence of y_d on x_j can be modeled by replacing x_{d0} with x_j in l_c .)

Figure 3.4 shows y_d versus V_{DS} predicted by (3.55) and by PISCES [Tec89] for 0.5- and 0.25- μm MOSFETs (described in Table 3.1) biased in weak inversion. For the numerical simulations, the boundary ($L - y_d$) is nebulous. In accord with our analysis, we defined it as the location along the channel where $d\phi_{fn}/dy \equiv -10E_y$. The numerical simulations support our model, while clearly refuting the commonly used 1-D model [Tay78] which predicts a relatively strong (square-root) V_{DS} dependence. The PISCES simulations do however indicate a slight dependence of y_d on V_{DS} . The discrepancy ($< 25\%$) shown in Fig. 3.4 diminishes as V_{DS} is reduced, suggesting the utility of our model for scaled technologies with supply voltages less than 5V. Some of this discrepancy is due to the nebulous nature of the modulated channel boundary, but we attribute most of it to the characterization of $\Delta E_x(x_{d0}, y)$, which in actuality is a component of a complex 2-D field. Our simple characterization in (3.52) however leads to the compact expression for y_d in (3.55), which does show accurate dependencies on technological parameters as verified in Fig.

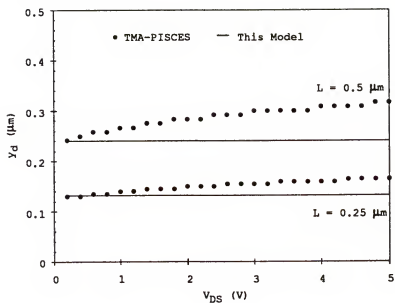


Fig. 3.4. Drain-side channel-length reduction versus drain-source bias in weak inversion predicted by PISCES and the model of (3.55). Device parameters are listed in Table 3.1.

TABLE 3.1
Device Parameters [Dor92] Used in the Simulations of Fig. 3.4

L (μm)	N_B (cm^{-3})	t_{ox} (nm)	x_j (μm)
0.50	1.5×10^{17}	13.3	0.20
0.25	4.0×10^{17}	7.6	0.12

3.4. In this regard our model is scalable and is superior to both the 1-D model of [Tay78], which only has dependence on body doping (N_B), and the empirical model [She87], which only has empirical dependence on L . Insight regarding device scaling is thus obvious from the physical yet simple model.

We extrapolate the model for y_d to $V_{DS} = 0$ (see Fig. 3.4) to evaluate the source-side modulation. The result is $y_s = y_d$, since near the source the quasi-Fermi-level separation (viz., voltage) is negligibly small. Therefore, we obtain the total modulated channel length in weak inversion,

$$L_e = L - y_s - y_d \cong L - 2l_c \ln\left(\frac{L^2}{l_c^2}\right) \equiv L - \Delta L. \quad (3.56)$$

Now in order to evaluate the channel current using (3.50) we need to evaluate (3.49) at the modulated channel boundaries y_s and $L - y_d$. Throughout the modulated channel the total surface potential is $\psi_s = \psi_{s0} + \Delta\psi_s$, where $\Delta\psi_s$ results from DIBL and is expressed by (3.24). In accord with Fig. 3.2(b), the surface potential when $V_{DS} = 0$, ψ_{s0} , is computed using (3.18). Hence ψ_s , which is not a function of y for $y_s \leq y \leq L - y_d$, is completely characterized. Therefore from (3.50) with (3.49)

$$I_{CH} = \frac{W}{L_e} \frac{q n_i^2 \bar{\mu}_n}{N_B} \frac{V_{th}^2}{E_x} \exp\left(\frac{\psi_s + V_{BS}}{V_{th}}\right) \left[1 - \exp\left(\frac{-V_D(L - y_d)}{V_{th}}\right) \right], \quad (3.57)$$

where $V_D(L - y_d)$ is the added separation of electron and hole QFL's at $y = L - y_d$ due to V_{DS} . This separation will be larger than a few V_{th} for most drain-source biases, so the exponential term will be negligible; it is convenient to let $V_D(L - y_d) = V_{DS}$. $V_D(y_s)$ is negligible. The average channel electron mobility, $\bar{\mu}_n$, is computed using (3.30) which accounts for transverse-field mobility degradation. This field is defined as $\bar{E}_x \triangleq \bar{E}_{x0} + \Delta E_x$, where \bar{E}_{x0} is the average field due to V_{GS} and V_{BS} with $V_{DS} = 0$, and ΔE_x represents the change in this field due to $V_{DS} > 0$. \bar{E}_{x0} is expressed using Gauss's law,

$$\bar{E}_{x0} = - \frac{Q_b}{\epsilon_s} \quad (3.58)$$

where Q_b is the body depletion-charge density for $V_{DS} = 0$ (3.3). The field ΔE_x is similarly derived

$$\Delta E_x = - \frac{\Delta Q_b}{\epsilon_s} , \quad (3.59)$$

and can be expressed using the results of the DIBL analysis. Gauss's law and (3.21) with the approximation $\Delta \Psi_s(y) \cong \Delta \Psi(x_{d0}, y)$ gives

$$\Delta Q_b = \frac{\epsilon_s x_{d0} \eta}{2} , \quad (3.60)$$

which expresses the reduction in charge that results from the increased charge-sharing caused by the drain bias. Hence the weak-inversion channel current is fully characterized.

For MOSFETs that are not properly scaled L_e in (3.56) can approach zero, and may become negative. These devices will suffer significant undesirable short-channel effects and/or punchthrough, and should be redesigned. Numerical simulations showed that MOSFETs with L_e much less than zero exhibit punchthrough. However MOSFETs with L_e near zero or slightly negative exhibit enhanced current relative to (3.57), even at low V_{DS} , with subthreshold slopes that differ negligibly from longer-channel devices. This model includes a semi-empirical modification to (3.57) to estimate the performance of the latter devices.

We limit L_e in the denominator of (3.57) to a minimum value of one extrinsic Debye length [Sze81], $L_{e(min)} = \sqrt{\epsilon_s v_{th} / (qN_B)}$. (Note this limiting value is consistent with the one rigorously derived by Taylor [Tay78] for the 1-D model.) Therefore the only dependence of I_{CH} on L is through the DIBL model via Ψ_s , which is negligible for low V_{DS} . To model the short-channel enhancement in current when L_e is limited and DIBL is negligible, \bar{E}_x is decreased in (3.57) in accordance with the physical mechanism of charge sharing, which decreases \bar{E}_x by reducing (in magnitude) the gate-controlled body depletion charge Q_b . We empirically model this effect by modifying Q_b in (3.58) when $(L - y_s - y_d) < L_{e(min)}$:

$$Q'_b = \frac{LQ_b}{Y_s + Y_d + L_{e(\min)}} \quad (3.61)$$

for which $|Q'_b| < |Q_b|$.

3.5 Cubic Spline for Moderate Inversion Operation

A cubic spline is used to interpolate between weak- and strong-inversion transport current characteristics, forming the moderate-inversion characteristic. The spline is applied to the logarithm of the channel current and ensures the continuity of the current and transconductances across the region transitions:

$$\ln(I_{CH}) = \alpha_0 + \alpha_1(V_{GS} - V_{TW}) + \alpha_2(V_{GS} - V_{TW})^2 + \alpha_3(V_{GS} - V_{TW})^3. \quad (3.62)$$

The coefficients in (3.62) are functions of V_{TW} and V_{TS} as well as the currents and transconductances at these voltages:

$$\alpha_0 = \ln[I_{CH}^{wk}(V_{TW})], \quad (3.63)$$

$$\alpha_1 = \frac{d \ln[I_{CH}^{wk}(V_{TW})]}{dV_{GS}}, \quad (3.64)$$

$$\alpha_2 = \frac{3 \ln[I_{CH}^{str}(V_{TS}) / I_{CH}^{wk}(V_{TW})]}{(V_{TS} - V_{TW})^2} - \frac{\frac{d \ln[I_{CH}^{str}(V_{TS})]}{dV_{GS}} + 2 \frac{d \ln[I_{CH}^{wk}(V_{TW})]}{dV_{GS}}}{(V_{TS} - V_{TW})}, \quad (3.65)$$

and

$$\alpha_3 = \frac{2 \ln[I_{CH}^{wk}(V_{TW}) / I_{CH}^{str}(V_{TS})]}{(V_{TS} - V_{TW})^3} + \frac{\frac{d \ln[I_{CH}^{str}(V_{TS})]}{dV_{GS}} + \frac{d \ln[I_{CH}^{wk}(V_{TW})]}{dV_{GS}}}{(V_{TS} - V_{TW})^2} \quad (3.66)$$

The derivative terms are computed analytically from the weak-inversion model and numerically using a difference approximation from the strong-inversion model.

3.6 Charging Currents of the Intrinsic Device

The source of charging currents is the time-dependent bias-induced modulation of space charge. These currents are expressed as $dQ/dt (= \sum_i (\partial Q / \partial V_i) (dV_i/dt))$ for i node-voltage differences), where Q is the total space charge associated with each terminal. These charges are computed from the preceding models for weak and strong inversion, and via linear interpolation in moderate inversion:

$$Q_X^{mod} = Q_X^{wk} + (Q_X^{str} - Q_X^{wk}) \frac{V_{GS} - V_{TW}}{V_{TS} - V_{TW}}, \quad (3.67)$$

where X is S (source), D (drain), and G (gate). In (3.67), V_{TW} and V_{TS} are the weak- and strong-inversion threshold voltages. The voltage derivatives of the charges are calculated using difference approximations. The derivations of the expressions

for the charges in strong inversion rely on the following mathematical manipulations [Vee88]. Consider a section $y_a \leq y \leq y_b$ along the channel. With (3.29) for $v(y)$ expressed using the spatially-constant mobility $\bar{\mu}_{(eff)}$ of (3.35), (3.27) can be written as

$$I_{DS} + \frac{I_{DS}\bar{\mu}_{(eff)}}{2v_{sat}} \frac{d\psi_s}{dy} = -WQ_c(y)\bar{\mu}_{(eff)} \frac{d\psi_s}{dy} . \quad (3.68)$$

With (3.22), (3.68) gives

$$dy = - \frac{\bar{\mu}_{(eff)}W}{I_{DS}(C_{ox} + \frac{\epsilon_s}{x_{d0}})} (Q_c + \frac{I_{DS}}{2v_{sat}W}) dQ_c . \quad (3.69)$$

To simplify the analysis we define

$$Q' \triangleq Q_c + \frac{I_{DS}}{2v_{sat}W} , \quad (3.70)$$

which combined with (3.69) produces

$$Q'(y) = Q'(y_a) \left\{ 1 - \left[1 - \frac{Q'^2(y_b)}{Q'^2(y_a)} \right] \frac{y - y_a}{y_b - y_a} \right\}^{1/2} \\ \triangleq Q'(y_a) \left[1 - K \frac{y}{y_{ba}} \right]^{1/2} . \quad (3.71)$$

These expressions are used to derive the following integrals from which the terminal charges are computed:

$$\int_{y_a}^{y_b} \Delta\psi_s dy = \frac{Q'(y_a) y_{ba}}{C_{ox} + \frac{\epsilon_s}{x_{d0}}} \left\{ \frac{2}{3K} [1 - (1 - K)^{3/2}] - 1 \right\}, \quad (3.72)$$

$$\int_{y_a}^{y_b} Q_c dy = -\frac{I_{DS} y_{ba}}{2v_{sat}W} + \frac{2Q'(y_a) y_{ba}}{3K} [1 - (1 - K)^{3/2}], \quad (3.73)$$

and

$$\begin{aligned} \int_{y_a}^{y_b} \frac{Y}{L} Q_c dy = & -\frac{I_{DS}(y_b^2 - y_a^2)}{4v_{sat}WL} + \frac{Q'(y_a) y_{ba}}{KL} \left\{ \frac{2}{3} [y_a - y_b(1 - K)^{3/2}] \right. \\ & \left. + \frac{4y_{ba}}{15K} [1 - (1 - K)^{5/2}] \right\}. \end{aligned} \quad (3.74)$$

Algebraic manipulation of the above equations with $y_a = 0$ and $y_b = L$ yields the charge expressions given below. Note that to simplify the analysis of the terminal charges we do not include the depletion charge due charge sharing; this is justified because these charges are much less than the depletion charge the drain, source, and body support in the extrinsic device regions.

3.6.1 Gate Charge

The gate charge is computed from

$$Q_G = WC_{ox} \int (V_{GS} - V_{BS} - \Phi_{ms} - \Psi_s) dy \quad (3.75)$$

where Φ_{ms} is the gate-semiconductor work-function difference. In weak inversion we simplify the analysis by accounting only for the charge within the region bound by $y_s \leq y \leq y_d$, i.e., the modulated channel length (L_e) given in (3.56). Therefore, (3.75) gives

$$Q_G^{wk} = WC_{ox} L_e [V_{GS} - V_{BS} - \Phi_{ms} - \Psi_s] \quad (3.76)$$

where Ψ_s is evaluated using (3.18) and (3.24).

In strong inversion the gate charge is computed from the integration of (3.75), with (3.72), from $y = 0$ to $y = L$:

$$Q_G^{str} = WC_{ox} L \left[V_{GS} - V_{BS} - \Phi_{ms} - \Psi_{ss} - \frac{V_{DS}}{2} + \frac{V_{DS}^2 (s + 1)}{12u - 6V_{DS}} \right] \quad (3.77)$$

where $s \triangleq \bar{\mu}_{eff} V_{DS} / 2V_{sat} L$ and $u \triangleq -Q_c(0) / (C_{ox} + \epsilon_s / x_{do})$. In saturation, L is replaced by L_e , and V_{DS} by $V_{DS(eff)}$ in (3.77), which is supplemented with

$$Q_G^S = WC_{ox} \{ (L - L_e) (V_{GS} - V_{BS} - \Phi_{ms} - \Psi_{ss} - V_{DS(eff)}) + \eta l_c^2 \} \\ - \frac{2V_{sat} l_c^2}{\bar{\mu}_{eff}} \left[\cosh\left(\frac{L - L_e}{l_c}\right) - 1 \right] - \eta l_c^3 \sinh\left(\frac{L - L_e}{l_c}\right) \} \quad (3.78)$$

to account for the charge in the high-field region. Equation (3.78) was derived by integrating (3.75) from $y = L_e$ to $y = L$, using the solution of (3.45) to express $\Delta\psi_s(y)$.

3.6.2 Drain Charge

In weak inversion $Q_D = 0$. In strong inversion a simple scheme for partitioning the channel charge between the source and drain [Oh80] is used for computing Q_D ,

$$Q_D = W \int_0^L \frac{y}{L} Q_c(y) dy$$

$$= -WL(C_{ox} + \frac{\epsilon_s}{x_{d0}})V_{DS} \left[\frac{2(z-1)^3}{3(2z-1)} + \frac{4[z^5 - (z-1)^5]}{15(2z-1)^2} + \frac{u-z}{2} \right] \quad (3.79)$$

where $z \triangleq u/V_{DS} - I_{CH}/2v_{sat}W(C_{ox} + \epsilon_s/x_{d0})V_{DS}$; to obtain (3.79) we used (3.74). In saturation L is replaced by L_e , and V_{DS} by $V_{DS(eff)}$ in (3.79). The supplemental high-field region charge is

$$Q_D^s = \frac{WQ_c(L_e) [L^2 - L_e^2]}{2L}, \quad (3.80)$$

which was derived with a partitioning factor of y/L , using a constant channel charge $Q_c(L_e)$, computed from (3.28) and (3.41), to describe the high-field region.

3.6.3 Source Charge

In weak inversion $Q_S = 0$. In strong inversion $Q_S = Q_{CH} - Q_D$ where the total channel charge is

$$Q_{CH} = W \int_0^L Q_c(y) dy$$

$$= -WL(C_{ox} + \frac{\epsilon_s}{x_{d0}}) V_{DS} \left[\frac{2}{3} \frac{z^3 - (z-1)^3}{2z-1} + u - z \right], \quad (3.81)$$

which was derived using (3.73). In saturation L is replaced by L_e , and V_{DS} by $V_{DS(eff)}$, and the supplemental high-field region charge is found from $WQ_c(L_e)[L - L_e] - Q_D^S$:

$$Q_S^S = WQ_c(L_e) \left[L - L_e - \frac{(L^2 - L_e^2)}{2L} \right]. \quad (3.82)$$

3.6.4 Body Charge

The body charge is computed using the condition of charge neutrality:

$$Q_B = -(Q_G + Q_D + Q_S + Q_f WL) \quad (3.83)$$

where Q_f is the fixed charge in the oxide.

3.7 Effects of LDD/LDS

The LDD and LDS regions affect the biases on the intrinsic structure used in the model. Due to the complexity

of the physics associated with the LDS and the LDD it is necessary to make simplifying assumptions regarding the structure of the device in order to maintain acceptable computational efficiency. Consequently this model assumes a symmetric non-overlapping LDD/LDS structure.

3.7.1 Voltage Drop Across the LDS

The voltage drop across the non-overlapped LDS is modeled simply by a resistor,

$$V_{LDS} = I_{CH} R_{LDS}, \quad (3.84)$$

so the biases V_{GS} , V_{DS} , and V_{BS} in the intrinsic device model of the previous sections are reduced by V_{LDS} . Assuming uniform current density across the LDS thickness, the resistance of the LDS, R_{LDS} , is computed from

$$\frac{1}{R_{LDS}} = \frac{W}{L_{LDS}} \int_0^{x_j(LDS)} \sigma(x) dx = \frac{W}{L_{LDS}} \bar{\mu}_n(LDS) q N_{LDS} x_j(LDS) \quad (3.85)$$

where the LDS is characterized by its length, L_{LDS} , doping density, N_{LDS} , junction depth, $x_j(LDS)$, and an average electron mobility, $\bar{\mu}_n(LDS)$.

3.7.2 Voltage Drop Across the LDD

The voltage drop across the LDD, V_{LDD} , is computed from the field distribution in the region [Cho91]. Figure 3.5

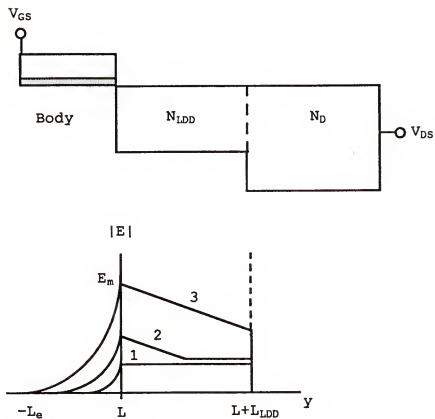


Fig. 3.5. The drain-side region of the LDD MOSFET, with possible longitudinal electric field distribution as a function of V_{DS} [Cho91].

shows the model drain-side structure, and possible longitudinal electric field distributions for different drain biases. Neutral regions are characterized by a constant field, which define simple ohmic drops. Depletion regions are induced in the LDD by a high channel field E_m , and are characterized by a graded electric field. Assuming the electron charge is negligible in the LDD, Poisson's equation describes the depletion region with a constant field slope, S :

$$\frac{d|E|}{dy} = \frac{-qN_{LDD}}{\epsilon_s} \equiv -S \quad (3.86)$$

where N_{LDD} is the doping density of the LDD.

The field at the body-LDD junction, E_m , can be computed from the expressions for $d\psi_s/dy$ (at $y = L$) obtained from the channel-length modulation analyses for weak and strong inversion. It can be easily shown that E_m from both these expressions can be accurately approximated by

$$E_m \equiv \frac{V_{DS} - V_{LDD} - V_{DS(eff)}}{l_c} \quad (3.87)$$

where V_{LDD} is the voltage drop across the LDD, and l_c is defined previously. In strong inversion $V_{DS(eff)}$ is computed in the channel analysis, and in weak inversion $V_{DS(eff)} \equiv 0$. In moderate inversion $V_{DS(eff)}$ and l_c are calculated via linear interpolation. The bias $(V_{DS} - V_{LDD})$ replaces V_{DS} in the intrinsic device model. Obviously then the characterization

of V_{LDD} imposes an iterative solution procedure. For an assumed value of V_{LDD} , the calculation of V_{LDD} is as follows.

For $(V_{DS} - V_{LDD}) \leq V_{DS(sat)}$, E_m is assumed small such that only the ohmic drop in the LDD is important, so field distribution '1' in Fig. 3.5 describes the LDD. Therefore,

$$V_{LDD} = I_{CH} R_{LDD}, \quad (3.88)$$

where

$$R_{LDD} = \frac{L_{LDD}}{q W N_{LDD} x_{j(LDD)} \bar{\mu}_n(LDD)} \quad (3.89)$$

In (3.89), the LDD is characterized by its length, L_{LDD} , doping density, N_{LDD} , junction depth, $x_{j(LDD)}$, and an average electron mobility, $\bar{\mu}_n(LDD)$. (Note these parameters are assumed equal to those that characterized the LDS in the previous subsection.)

For $(V_{DS} - V_{LDD}) > V_{DS(sat)}$, all three cases '1', '2', and '3' are possible. If $E_m \leq I_{CH} R_{LDD} / L_{LDD}$ then case '1' applies, and (3.88) is used to compute V_{LDD} . If $E_m \geq I_{CH} R_{LDD} / L_{LDD}$, either '2' or '3' can exist. The transition between '2' and '3' is defined by the condition $(E_m - I_{CH} R_{LDD} / L_{LDD}) = S L_{LDD}$, which can be derived from Fig. 3.5. For $(E_m - I_{CH} R_{LDD} / L_{LDD}) \leq S L_{LDD}$, condition '2' exists, and V_{LDD} is calculated using

$$V_{LDD} = \frac{(E_m - I_{CH} R_{LDD} / L_{LDD})^2}{2S} + I_{CH} R_{LDD} \quad (3.90)$$

For $(E_m - I_{CH} R_{LDD} / L_{LDD}) > S L_{LDD}$, condition '3' exists, and

$$V_{LDD} = 0.5 L_{LDD} (2E_m - S L_{LDD}). \quad (3.91)$$

3.8 Generation Current Due to Impact Ionization

In the saturation region ($V_{DS} > V_{DS(sat)}$), the model accounts for generation current due to (weak) impact ionization involving electrons traversing the high-field region near the drain:

$$I_{G1} = (M - 1) I_{CH}. \quad (3.92)$$

A local-field model [Slo87] is assumed for the ionization rate:

$$(M - 1) = \int \alpha_0 e^{-\beta_0 / E} dy \quad (3.93)$$

where α_0 and β_0 are both constant. With an LDD, impact-ionization current is generated both in the channel and in the LDD. In the former the multiplication factor in (3.93) can be evaluated as [Vee88]

$$(M - 1)_{CH} \equiv \frac{\alpha_0}{\beta_0} (V_{DS} - V_{DS(eff)}) e^{-\beta_0 l_c / (V_{DS} - V_{DS(eff)})} \quad (3.94)$$

which was derived using the expression for the peak channel electric field in (3.87); l_c and $V_{DS(eff)}$ were defined

previously for both weak and strong inversion, and are computed via linear interpolation in moderate inversion. With an LDD (and LDS) V_{DS} in (3.94) is reduced by the voltage drop(s) described in Section 3.7, hence a reduction of the generation current in the channel. The generation current in the LDD is evaluated by numerical (Romberg [Pre89]) integration of the linear field distribution in the region characterized in Section 3.7 [Cho91]:

$$(M - 1)_{LDD} \equiv \frac{\alpha_2}{S} \int_{E_m - SL_{LDD}}^{E_m} \exp\left(\frac{-\beta_0}{|E|}\right) d|E| \quad (3.95)$$

where the lower limit of integration is set to zero if $(E_m - SL_{LDD})$ is less than zero. The sum of (3.94) and (3.95) is the total multiplication factor used to compute I_{G1} in (3.92).

3.9 Summary

A model was derived for the short-channel LDD/LDS MOSFET which primarily relies on structure-dependent parameters, with minimal reliance on empirical modeling. The next chapter addresses model parameter evaluation and the implementation of the model in SUMM/MMSPICE. Chapter 4 also presents the extrinsic device parasitics which are modeled using conventional SPICE approaches [Ant88]. Those elements include source and drain series resistances (due to regions separate from the LDS/LDD), junction leakage diodes, and gate-to-source, gate-to-drain, and gate-to-body overlap capacitances.

CHAPTER 4 MODEL IMPLEMENTATION IN SUMM/MMSPICE

4.1 Introduction

This chapter describes the equivalent circuit of the charge-based MOSFET model developed in Ch. 3 (Fig. 4.1) and its implementation in the circuit simulator SPICE2 [Nag75]. This chapter also details the modifications of SUMM [Gre90] that were made to accommodate the MOSFET model, which include the addition of a new routine for evaluating several key model parameters. Some PISCES [Tec89] simulations that verify the evaluation algorithms and the MOSFET model are also presented.

The new model was implemented in the enhanced version of SPICE2, MMSPICE [Jeo90], which contains the physical charge-based bipolar model [Jeo89] in addition to all the standard SPICE2 device models. Implementation was accomplished by creating a new device type for our model, rather than implementing it as an additional level of the standard SPICE2 MOSFET device type. This approach was taken because of substantive differences between the formalism our charge-based model and that of the capacitance-based SPICE2 models.

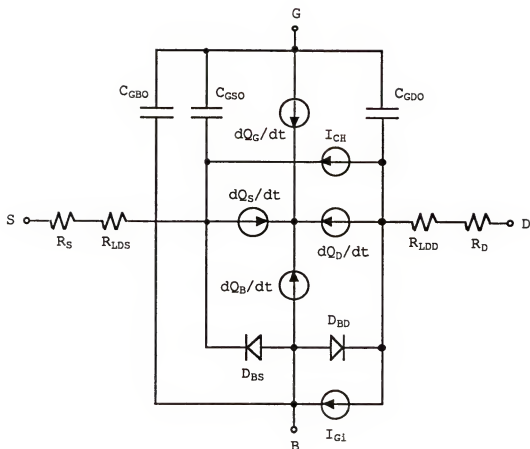


Fig. 4.1. The complete equivalent circuit representation of the quasi-static large-signal transient model for the bulk (n-)MOSFET.

4.2 Description of Equivalent Circuit Elements

The equivalent circuit shown in Fig. 4.1 contains elements that model the intrinsic and extrinsic regions of the MOSFET depicted in Fig. 3.1. The intrinsic region elements are I_{CH} , I_{G1} , dQ_G/dt , dQ_D/dt , dQ_S/dt and dQ_B/dt , which were all detailed in the derivation of the model in Ch. 3. The resistances R_{LDS} and R_{LDD} are assumed equal, and account for the ohmic drops across the LDS and LDD. The non-ohmic voltage drop across the LDD is accounted for in the element I_{CH} . R_{LDS} and R_{LDD} are computed in the MMSPICE subroutine MODCHK from model parameters (see (3.85) and (3.89)). Additional extrinsic resistances, e.g., contact resistance, are accounted for with the model parameters R_s and R_D .

The overlap capacitors C_{GBO} , C_{GDO} and C_{GSO} , modeled as parallel-plate capacitors, contribute charging currents, as do the junction diodes D_{BS} and D_{BD} . The voltage dependence of the diode depletion capacitance is modeled

$$C_D = CJO \left(1 - \frac{V_D}{PJ}\right)^{-MJ} \quad \text{for } V_D < \frac{PJ}{2} \quad (4.1)$$

and

$$C_D = CJO * 2^{MJ} \left[1 + MJ \left(\frac{2V_D}{PJ} - 1\right)\right] \quad \text{for } V_D \geq \frac{PJ}{2} \quad (4.2)$$

using standard SPICE representation. The model parameters in (4.1) and (4.2) are the zero-bias depletion capacitance, CJO ,

junction grading coefficient, MJ, and built-in potential, PJ. These diodes also contribute generation and conduction currents which are modeled using the conventional diode equation [Sze81] with model parameters JS, the saturation current density, and N, the junction ideality factor. The implementation of these elements in MMSPICE follows standard SPICE approaches [Ant88].

4.3 Implementation in MMSPICE

The equivalent circuit model was implemented by writing four new subroutines, QBMOD, QBFET, QBID, and QBEC, and modifying 17 of the original SPICE2 subroutines. This new device type is referenced by the letter "S", and the NMOS and PMOS versions of the model are referenced on the model call-line by QBNMOD and QBPMOD, respectively.

QBFET is similar to the subroutine MOSFET of the original SPICE2 source code. This routine calls the other three new routines as it computes and assembles the matrix elements (Y and RHS), i.e., conductances and currents, representing the circuit in Fig. 4.1 for dc and transient simulations. In QBID the model routine QBMOD is called to compute the elements I_{CH} and I_{G1} , their associated conductances and transconductances, and the capacitances and transcapacitances associated with the intrinsic-region elements dQ_G/dt , dQ_D/dt , dQ_S/dt and dQ_B/dt . These conductances and capacitances are calculated using finite difference approximations, necessitating multiple calls of QBMOD.

Any non-ohmic voltage drop across the LDD to support space (depletion) charge in that region is also accounted for in QBMOD. The detailed flow diagram of this subroutine is shown in Fig. 4.2; the non-ohmic voltage drop across the LDD (see Sec. 3.7.2) is represented by V_{LDD} . Iteration control is accomplished via a secant method assisted by bracketing and bisection [Pre89]. By directly placing R_{LDS} and R_{LDD} into the admittance matrix (in subroutine ACLOAD for ac simulations, and QBFET for dc and transients) MMSFICE can account for the voltage drops across these resistors due to all relevant currents associated with the overlap capacitors, junction diodes and the intrinsic device region. This implementation scheme is an improvement on the approach taken by Choi [Cho91] in which the total (ohmic and non-ohmic) voltage drops across the LDD and LDS are computed in the model subroutine, thereby accounting only for intrinsic-region transport current. Additionally, the simulation convergence speed for the new approach is superior, because it eliminates redundancy in the computations of parasitic ohmic voltage drops.

In QBEC the equivalent conductances of the intrinsic-region capacitors, the overlap capacitors and the junction diode capacitors are computed. The diode currents and the equivalent currents associated with all charging elements are computed in QBFET.

We estimate that transient simulations with our model will require three to four times the computational time of

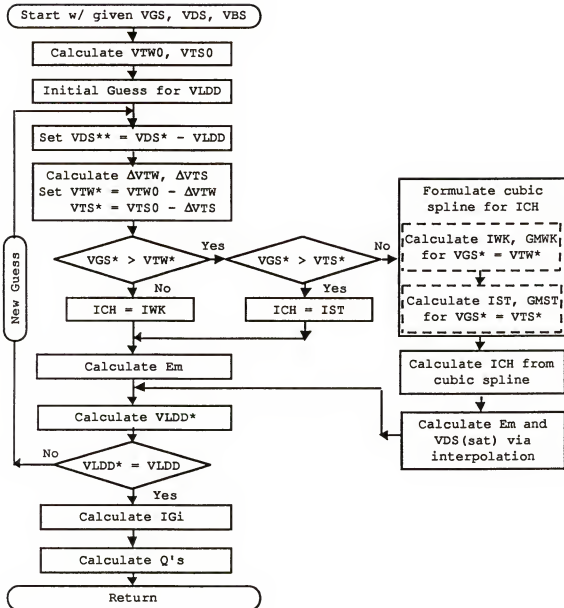


Fig. 4.2. The algorithm used in QBMOD for the model of the intrinsic device and the non-ohmic LDD voltage drop. The ohmic resistances of the LDD and LDS are accounted for in the subroutine QBFET.

simulations using the SPICE Level-2 MOSFET model [Ant88]. We attribute most of this difference to the multiple calls of our model routine QBMOD that are required to calculate the aforementioned conductances and capacitances.

4.4 Device and Model Parameters

The model is based on four device-line parameters (see Table 4.1) and 27 model parameters (see Table 4.2). The device-line parameters are the length (L) and width (W) of the gate, and the areas of the drain (AD) and source (AS). Of the 27 model parameters there are nine physical parameters which must be specified by the user. These are DL and DW, the channel length and width reductions which are due, for example, to lateral diffusion of the drain and source regions, the type of gate material, TPG, the gate oxide thickness, TOX, the fixed charge density (normalized by q) in the oxide, NQF, the source/drain junction depth, XJ, the length of the LDD/LDS, LLD, the doping of the LDD/LDS, NLD, and the (effective) body doping, NB. Rather than specifying the value of this latter parameter, the user can invoke SUMM to evaluate it from structural information.

Based on physical criteria, several key parameters can be evaluated by SUMM or by the routine MODCHK in MMSpice. From a user-supplied 1-D doping density profile under the gate SUMM evaluates an effective (constant) body doping density, NB, the low-field carrier mobility in the channel, UO, the body-drain/source junction built-in potential, PJ,

Table 4.1
MOSFET DEVICE-LINE PARAMETERS

<u>Name</u>	<u>Description</u>	<u>Units</u>
L	gate length	m
W	gate width	m
AD	drain area	m ²
AS	source area	m ²

Table 4.2
MOSFET MODEL PARAMETERS

Name	Description	Units
DL	channel-length reduction	m
DW	channel-width reduction	m
TOX	gate oxide thickness	m
NB	body doping	cm ⁻³
XJ	junction depth of lightly-doped drain/source (LDD/LDS) or heavily-doped drain/source (for LLD=0)	m
LLD	length of LDD/LDS	m
NLD	doping density of LDD/LDS	cm ⁻³
NQF	density of fixed charge in the gate oxide	cm ⁻²
TPG	type of gate material	-
	+1) opposite to body	
	-1) same as body	
	0) aluminum	
WKF	gate-body work-function difference	V
VFB	flatband voltage	V
U0	zero-field channel mobility	cm ² /Vs
THETA	mobility degradation coefficient	cm/V
VSAT	saturated carrier velocity	cm/s
ULD	average mobility in LDD/LDS	cm ² /Vs
AI	pre-exponential constant for impact-ionization coefficient	cm ⁻¹
BI	exponential constant for impact-ionization coefficient	V/cm

Table 4.2 -- continued

Name	Description	Units
JS	source- and drain-body diode saturation current density	A/m ²
N	source- and drain-body diode ideality factor	-
CJ0	zero-bias source- and drain-body junction capacitance	F/m ²
MJ	source- and drain-body junction exponential factor	-
PJ	source- and drain-body junction potential	V
CGD0	gate-drain overlap capacitance	F/m
CGS0	gate-source overlap capacitance	F/m
CGB0	gate-body overlap capacitance	F/m
RS	source parasitic resistance	Ω
RD	drain parasitic resistance	Ω

the zero-bias body-drain/source junction depletion capacitance C_{JO} , and the body-drain/source junction saturation current density, J_S . The routine MODCHK can evaluate an average mobility in the LDD/LDS, ULD, from NLD, the gate-body work-function difference, W_{KF} , and the flatband voltage, V_{FB} . The evaluation of the latter two parameters by MODCHK relies in part on the value of N_B computed by SUMM. The parameter evaluation routines are detailed in the next subsection.

There are two semi-physical parameters that must be specified by the user, the junction diode ideality factor, N , and exponential factor, MJ . There are five parameters modeling extrinsic parasitics that cannot be theoretically evaluated without equivocation, $CGSO$, $CGDO$, $CGBO$, RD and RS , which also must be specified by the user. There is one material constant, $VSAT$, the carrier saturation velocity, whose typical value is around 10^7 cm/s [Sze81]. Lastly, the model relies on only three empirical parameters: two of these are AI and BI , which are used to compute impact ionization current; typical values for AI and BI are 2.45×10^6 cm⁻¹ and 1.75×10^6 V/cm [Slo87], respectively, for n-MOSFETs. The third empirical parameter is $THETA$, which models the transverse-field dependence of the carrier mobility in the channel; typically its value is around 10^{-6} cm/V [Mul86].

4.5 Model Parameter Evaluation By SUMM and MODCHK

All parameters that are not specified by the user, i.e., those not included on the model card of the MMSPICE input file, are assigned default values in the MMSPICE routine MODCHK. There are seven parameters that can be evaluated by SUMM, together with MODCHK, based on analyses of a tabulated 1-D body (well) doping-density profile ($N_B(x)$) under the gate. The evaluation methodology preserves physical parametric correlations, and allows for a nonuniformly doped body which could result, for example, from a threshold-adjust (channel) implant; however, the evaluation is restricted to surface-channel MOSFETs. The user-specified profile, for which $x = 0$ corresponds to the silicon surface, must be of the following format: n-type doping densities must have a positive sign, while p-type densities must be negative; the units of depth (x) must be in microns and those of concentration in atoms/cm³. These four constraints are consistent with the format of profiles generated by SUPREM-3 [Tec88]. The execution of SUMM is illustrated in Sec. 4.6.

The evaluation of NB involves computing an effective (constant) body doping that optimally represents a nonuniformly doped body. The criterion for evaluating NB is based on accurately predicting threshold voltage and the corresponding surface potential Ψ_{ST0} . Using physical insight we define this potential ($V_{DS} = V_{BS} = 0$) for a MOSFET with nonuniform body doping as

$$\Psi_{sT0} \triangleq \phi_B(0) + \phi_B(\infty) \quad (4.3)$$

where

$$\phi_B(0) = \pm V_{th} \ln \left[\frac{N_B(0)}{n_i} \right] \quad (4.4)$$

is the potential due to the surface ($x = 0$) doping. This potential is positive for p-type surface doping and negative for n-type. In (4.3) $\phi_B(\infty)$ is the (reference) potential difference between the Fermi level and the intrinsic level of the well, and is evaluated in SUMM using the doping density at the deepest location in the well specified in the user's profile

$$\phi_B(\infty) = \pm V_{th} \ln \left[\frac{N_B(\infty)}{n_i} \right]. \quad (4.5)$$

This potential is positive for a p-well and negative for a n-well. For a uniformly doped body $\phi_B(0) = \phi_B(\infty)$, so $\Psi_{sT0} = 2\phi_B(0)$, the conventionally defined threshold surface potential with $V_{BS} = 0$.

A simple expression for the threshold voltage is derived from the summation of potentials and voltages ($V_{DS} = 0$) across the MOSFET. This summation for the actual device yields

$$V_{GS} - V_{BS} = \Phi_{gate} - \left[\chi_{Si} + \frac{E_{G(Si)}}{2q} + \phi_B(\infty) \right]$$

$$- \frac{Q_{s0} + Q_f}{C_{ox}} + \psi_{s0} \quad (4.6)$$

where Φ_{gate} is the work-function potential of the gate material, χ_{Si} is the electron affinity in silicon, $E_{G(Si)}$ is the silicon energy-band gap and the term in brackets is the body work-function potential. In (4.6) the potential drop across the oxide has been written in terms of charge in the silicon (Q_{s0}), fixed charge in the oxide (Q_f), and oxide capacitance (C_{ox}) using Gauss's law.

The expression for threshold voltage (for $V_{BS} = 0$) is obtained from (4.6) with $\psi_{s0} = \psi_{sT0}$,

$$V_{T0} \triangleq \Phi_{gate} - \chi_{Si} - \frac{E_{G(Si)}}{2q} - \frac{Q_{b0} + Q_f}{C_{ox}} + \phi_B(0) \quad (4.7)$$

where we have assumed Q_{s0} is predominantly depletion charge at this bias. Note that V_{T0} is used exclusively by SUMM to evaluate NB; it is not used in the MOSFET model. The actual threshold voltage of the device may differ from V_{T0} due to short-channel effects which are accounted for in the model as described in Chapter 3.

The expression from which NB is evaluated is derived by equating (4.7) to the corresponding expression of the conventionally defined threshold voltage for a MOSFET with uniform body doping NB:

$$V_{T0}' \equiv \Phi_{\text{gate}} - \left[\chi_{\text{sl}} + \frac{E_{\text{G}}(\text{sl})}{2q} + \phi_{\text{B}}(\text{NB}) \right] - \frac{Q_{\text{b0}}(\text{NB}) + Q_{\text{f}}}{C_{\text{ox}}} + 2\phi_{\text{B}}(\text{NB}) \quad (4.8)$$

where the gate-body work-function difference is the term in brackets, and the potential $\phi_{\text{B}}(\text{NB})$ is

$$\phi_{\text{B}}(\text{NB}) = \pm V_{\text{th}} \ln \left[\frac{\text{NB}}{n_{\text{I}}} \right]. \quad (4.9)$$

Equating (4.8) to (4.7) we obtain

$$-\frac{Q_{\text{b0}}}{C_{\text{ox}}} + \phi_{\text{B}}(\infty) \equiv -\frac{Q_{\text{b0}}(\text{NB})}{C_{\text{ox}}} + \phi_{\text{B}}(\text{NB}) \quad (4.10)$$

which governs the evaluation of NB. Terms on the right-hand side of (4.10) are analytic functions of this parameter, while the left-hand side is a function of the actual body doping profile $N_{\text{B}}(x)$. Note that both sides depend on the value of the model parameter TOX ($= \epsilon_{\text{ox}}/C_{\text{ox}}$), indicating an important correlation between this parameter and NB in our evaluation. The depletion charge density $Q_{\text{b0}}(\text{NB})$ in (4.10) can be expressed analytically by solving the 1-D Poisson equation for uniform body doping, i.e., (3.3). This charge density is a function of the potential $\psi_{\text{s}} = 2\phi_{\text{B}}(\text{NB})$, where $\phi_{\text{B}}(\text{NB})$ is given in (4.9). $\phi_{\text{B}}(\infty)$ is computed using (4.5).

Numerical integration of $N_{\text{B}}(x)$ is required to find the depletion charge density Q_{b0} (for $\psi_{\text{s0}} = \psi_{\text{sT0}}$) in (4.10):

$$Q_{b0} = \pm q \int_0^{x_{d0}} N_B(x) dx . \quad (4.11)$$

The depletion-layer depth x_{d0} in (4.11) is calculated via numerical integration of the 1-D Poisson equation (3.1). Using the method of integration-by-parts on (3.1), with the depletion approximation (i.e., $E_x(x_{d0}) \equiv 0$), we derive

$$\psi_{s0} - \psi(x_{d0}) = \frac{q}{\epsilon_s} \int_0^{x_{d0}} x N_B(x) dx . \quad (4.12)$$

The potential at x_{d0} in (4.12) is

$$\psi(x_{d0}) = \phi_B(\infty) - \phi_B(x_{d0}) \quad (4.13)$$

for $V_{BS} = 0$, where

$$\phi_B(x_{d0}) = \pm V_{th} \ln \left[\frac{N_B(x_{d0})}{n_i} \right] ; \quad (4.14)$$

this potential is positive for p-type doping and negative for n-type. Notice that (4.12) implicitly defines x_{d0} . To compute this depth using (4.12) we impose the boundary condition $\psi_{s0} = \psi_{s00}$, then iteratively search for a value of x_{d0} which satisfies that equation, computing its right-hand side using numerical integration. After finding x_{d0} , Q_{b0} is computed using (4.11). Then with (3.3) and (4.9) N_B is evaluated from (4.10) via Newton-Raphson iteration [Pre89]. For the trivial

case of uniform body doping SUMM returns the constant density value.

The low-field carrier mobility in the channel, μ_0 , is evaluated from $N_B(x)$ using measurement-based empirical relations [Mul86] for majority-carrier bulk mobility. For n-MOSFETs we use

$$\mu_n = 52.2 + \frac{1364.8}{1 + \left(\frac{N}{9.68 \times 10^{16}} \right)^{0.68}} \quad (4.15)$$

and for p-MOSFETs

$$\mu_p = 44.9 + \frac{425.6}{1 + \left(\frac{N}{9.2 \times 10^{16}} \right)^{0.711}} \quad (4.16)$$

The surface doping density is used to evaluate (4.15) and (4.16), i.e., $N = N_B(0)$. We acknowledge that using these equations for channel mobility neglects surface scattering effects, however, this is not unreasonable because of the high body doping density in scaled devices.

To evaluate the parameters PJ , CJO and JS , which characterize the extrinsic source-body and drain-body junctions, SUMM requires a representative value for the extrinsic-region body doping density $N_{B(\text{ext})}$. By default SUMM approximates $N_{B(\text{ext})}$ with the intrinsic-region doping density at the junction depth XJ , $N_B(XJ)$, where XJ is a user-specified parameter. SUMM prompts the user for an optional alternative

value for $N_{B(\text{ext})}$ that may be more accurate than the one assumed.

With $N_{B(\text{ext})}$ the built-in junction potential PJ is calculated:

$$PJ = \frac{E_G(\text{Si})}{2q} + V_{th} \ln \left[\frac{N_{B(\text{ext})}}{n_i} \right], \quad (4.17)$$

assuming high doping densities in the source and drain.

The zero-bias junction capacitance CJO is evaluated using a one-sided abrupt-junction approximation:

$$CJO = \sqrt{\frac{q \epsilon_s N_{B(\text{ext})}}{2PJ}}. \quad (4.18)$$

SUMM computes the saturation current density JS (written for an n-MOSFET) as

$$JS = q \sqrt{\frac{D_n}{\tau_n}} \frac{n_i^2}{N_{B(\text{ext})}} \quad (4.19)$$

where $D_n = v_{th} \mu_n$ is the minority-electron diffusivity and τ_n is the minority-electron lifetime in the body. The corresponding JS for a p-MOSFET is calculated using D_p and τ_p . Carrier mobilities in the body are computed in SUMM using (4.15) for n-MOSFETs and (4.16) for p-MOSFETs, with $N = N_{B(\text{ext})}$. The carrier lifetimes are calculated from empirical expressions obtained from measurements [Sol90]:

$$\tau_n = \frac{3.0 \times 10^{-5}}{1 + \frac{N}{1.0 \times 10^{17}}} \quad (4.20)$$

for n-MOSFETs, and

$$\tau_p = \frac{1.0 \times 10^{-5}}{1 + \frac{N}{1.0 \times 10^{17}}} \quad (4.21)$$

for p-MOSFETs, with $N = N_{B(\text{ext})}$.

Three model parameters, WKF, VFB and ULD, are evaluated in the MMSPICE routine MODCHK, using values of other model parameters. The gate-body work-function difference WKF is computed

$$\text{WKF} = \Phi_{\text{gate}} - \Phi_{\text{body}}, \quad (4.22)$$

where Φ_{gate} and Φ_{body} are the work-function potentials of the gate and body, respectively. For a polysilicon gate, Φ_{gate} is computed from material constants and user-specified parameters:

$$\Phi_{\text{gate}} = \chi_{\text{Si}} + \frac{E_{\text{g}}(\text{Si})}{2q} \pm \text{TPG} \frac{E_{\text{g}}(\text{Si})}{2q} . \quad (4.23)$$

For this evaluation we use χ_{Si} equal to 4.15 V at 300 K [Sze81], and $E_{\text{g}}(\text{Si})$ equal to 1.12 eV at 300 K [Sze81]. The type of MOSFET determines which sign is used before the last term in (4.23): the minus sign is used for p-MOSFETs, and the plus

sign for n-MOSFETs; the parameter TPG is (+1) if the body and gate are of opposite doping type and (-1) if they are the same type. For TPG = 0, an aluminum gate is assumed with $\Phi_{\text{gate}} = 4.1 \text{ V}$ [Sze81].

The body work-function is obtained from (4.8):

$$\Phi_{\text{body}} = \chi_{\text{Si}} + \frac{E_{\text{g}}(\text{Si})}{2q} + \phi_{\text{B}}(\text{NB}) , \quad (4.24)$$

using (4.9). The work-function difference in (4.22) is thus obtained.

The gate flatband voltage VFB is evaluated in MODCHK using WKF, NQF and TOX:

$$\text{VFB} = \text{WKF} - \frac{\text{TOX} * q * \text{NQF}}{\epsilon_{\text{ox}}} . \quad (4.25)$$

The average mobility in the LDD and LDS, ULD, is computed in MODCHK using (4.15) for NMOS and (4.16) for PMOS, with $N = \text{NLD}$. This mobility is used in the model for calculating the resistances of these regions.

4.6 Implementation in SUMM

This section presents the changes made to the program SUMM [Gre90] to implement the parameter evaluation described in Sec. 4.5. These changes include modifications to the Bourne shell script in the file summ that performs the menu-driven user interface, modifications to four of the original

SUMM files (written in C), head.h, main.c, read_profile.c and write_output.c, and the introduction of an additional routine params5 that evaluates the MOSFET parameters.

The new main menu is depicted in Fig. 4.3. It differs slightly from the original: model parameter evaluations and MMSPICE execution are now separate options. Options 1 and 3 are handled in procedures similar to the original SUMM, and therefore are not described here.

Selection of option 2 leads the new "Device Model Selection Menu" shown in Fig. 4.4. From this menu the user chooses which charge-based MMSPICE model card to edit. Upon selection, the user is prompted for the name of the MMSPICE input file that contains circuit information, the (required) model call-line, and any user-specified model parameters. Selection of 1 from this menu is detailed in [Gre90]. The procedures following selections 2 and 3 for NMOS and PMOS, respectively, are identical, so only one of these is described below.

Choosing option 2 of the device model selection menu leads to the "NMOS Model Card Parameter Editing Menu", shown in Fig. 4.5. Displayed above the menu options are the model call-line and parameters contained in the user's MMSPICE input file. Any parameters not specified are evaluated in SUMM or MODCHK (where they may simply be assigned default values.) Details of options 1 and 3 are contained in [Gre90]. Option 2 enables the user to evaluate the unspecified

SUMM: A SUPREM-3 / MMSPICE-3 Integrator

- 1) SUPREM-3
- 2) Parameter Evaluation
- 3) MMSPICE-3
- 4) Exit SUMM

(Bourne shell commands can be executed from
the response prompt)

Selection:

Fig. 4.3. The main menu of the parameter evaluator SUMM.

Device Model Selection Menu

- 1) QBBJT (npn bipolar)
- 2) QBNMOS (n-MOSFET)
- 3) QBPMOS (p-MOSFET)
- 4) Quit to Main Menu

(Bourne shell commands can be executed from
the response prompt)

Selection:

Fig. 4.4. At this menu the user selects a charge-based
MMSPICE device model for parameter evaluation.

```
.MODEL model_name QBNMOS
+ parameter_name=parameter_value...
.
.
```

NMOS Model Card Editing Menu

- 1) Change the Value of an Individual Parameter
- 2) Evaluate the Following Parameter(s) From a Doping Profile:
NB UO PJ CJO JS
- 3) Create a New MMSPICE Input File Comprising the Displayed Model Card
- 4) Quit to Device Model Selection Menu

(Bourne shell commands can be executed from the response prompt)

Selection:

Fig. 4.5. The NMOS model card editing menu. Above the menu is a display of the model call-line and any parameters contained in the user's MMSPICE input file.

parameters NB, UO, PJ, CJO and JS from a tabulated 1-D body doping density profile under the gate.

Figure 4.6 depicts how the file containing the tabulated profile is entered. To evaluate NB SUMM requires a value of the parameter TOX. If the user has not specified a value in the input circuit file SUMM assumes $\text{TOX} = 15 \text{ nm}$. To evaluate PJ, CJO, and JS, SUMM requires a value of XJ (default = $0.2 \mu\text{m}$) to approximate the extrinsic body doping density at this junction depth using the tabulated profile; or, the user may enter a more representative value for this density at the response prompt also depicted in Fig. 4.6.

4.7 PISCES Support

Figure 4.7 shows two doping-density profiles that are representative of MOSFET body doping before and after a threshold-adjust (channel) implant. Included on the figure is the effective doping density computed by SUMM from the nonuniform profile. The MMSPICE input file with the SUMM-assembled model card for the implanted-channel device is shown in Fig. 4.8. SUMM/MMSPICE and PISCES [Tec89] predictions of I_{DS} versus V_{GS} characteristics ($V_{\text{DS}} = 0.05 \text{ V}$, $V_{\text{BS}} = 0.0 \text{ V}$) using both the implanted and unimplanted profiles are compared in Fig. 4.9. The technology parameters used for the simulations were $L = 0.8 \mu\text{m}$, $W = 20 \mu\text{m}$, $t_{\text{ox}} = 20 \text{ nm}$ and $x_j = 0.3 \mu\text{m}$. The predicted characteristics from both tools are quite similar, and serve to verify the predictive capability of the new model with the SUMM parameter evaluation. No

```
.MODEL model_name QBNMOS  
+ parameter_name=parameter_value...  
.  
.  
.
```

Enter the tabulated profile file name: *filename*

Enter a value of extrinsic body doping density for the evaluation of source-body and drain-body junction parameters [default=density (atoms/cm³) in *filename* at junction depth XJ]:

Fig. 4.6. Selection of option 2 from the NMOS model card editing menu leads to the parameter evaluation detailed in Sec. 4.5.

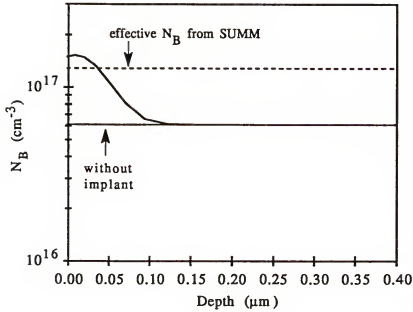


Fig. 4.7. The body doping profiles for an n-MOSFET. The nonuniform profile depicts a representative channel doping implant for threshold voltage adjustment. The SUMM-evaluated effective doping density for the nonuniform profile is also shown. For the evaluation $t_{ox} = 20 \text{ nm}$.


```

THIS IS AN EXAMPLE MMSPICE INPUT FILE FOR MOSFET SIMULATION
S1 1 2 0 0 QBNMOD W=20U L=0.8U AS=30P AD=30P
VDS 2 0 0.05
VGS 1 0 2.5
.MODEL QBNMOD QBNMOS
+ TOX=20.0E-9 XJ=0.3E-6 NB=1.26E17 UO=639
+ PJ=0.96 CJO=7.28E-4 JS=5.10E-9 THETA=1.0E-6
.DC VGS 0 5 0.1
.PRINT DC I(VDS)
.END

```

Fig. 4.8. An example of an MMSPICE input file that executes the new MOSFET model. The model card contains the SUMM-evaluated effective body doping density NB, as well as the (correlated) parameters UO, PJ, CJO and JS, for the nonuniform profile shown in Fig. 4.7.

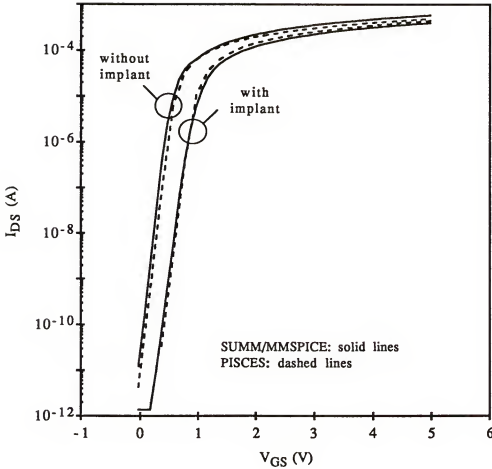


Fig. 4.9. Comparisons of PISCES [Tec89] and SUMM/MMSPICE predictions of I_{DS} versus V_{GS} characteristics ($V_{DS} = 0.05$ V, $V_{BS} = 0.0$ V) for a n-MOSFET before and after a representative threshold-adjust implant. Body doping profiles are shown in Fig. 4.7. For the simulations $L = 0.8$ μm , $W = 20$ μm , $t_{ox} = 20$ nm and $x_j = 0.3$ μm .

matching of mobility models was done, which accounts for some of the discrepancy apparent in these characteristics. Some discrepancy is also attributed to differences in simulation formalism, i.e., the seminumerical simulator based on quasi-2-D analyses is not as accurate as the fully-numerical 2-D analysis; however, the MMSPICE simulations were obtained in a small fraction of the time required for the PISCES simulations. Indeed, as seen in Fig. 4.9 the predictions of threshold voltage shift due to the implant, which are reflected by the shifts in subthreshold characteristics, are nearly equal.

CHAPTER 5 MODEL VERIFICATION/APPLICATION

5.1 Introduction

This chapter presents some comparisons between MOSFET model simulations and actual device and circuit measurements. These comparisons provide support for the pragmatic MOSFET device/circuit simulator. In addition the measurements indicate limitations inherent in this and all existing circuit-based short-channel MOSFET models, suggesting the need for additional, innovative modeling efforts. Based on the simulations of the test devices, which were fabricated in a merged MOSFET/BJT (BiCMOS) technology, the SUMM/MMSPICE tool is applied to suggest how the technology could be optimally scaled from the 0.6- μm effective channel lengths down to 0.4 μm (n-MOSFETs) and 0.35 μm (p-MOSFETs).

5.2 Comparisons with Device and Circuit Measurements

For these comparisons we make use of a Semiconductor Research Consortium member company's 0.8 μm dual-polysilicon twin-well LDD/LDS CMOS technology. In order to demonstrate the utility of our model as a device simulator we evaluate key model parameters based on some knowledge of the technology in concert with the evaluation method described in Ch. 4. In addition, several parameters are evaluated or tuned

from the measured characteristics; however, little effort is given to generating the "best fit" to the measured data.

From the device manufacturer we learned that the effective (metallurgical) channel lengths of these devices are $0.6\text{ }\mu\text{m}$, and the gate oxide thickness for the technology is 15 nm . The devices compared in this section have effective channel widths of $19\text{ }\mu\text{m}$. The technology utilizes high well (body) doping densities to minimize short-channel effects and avoid source-drain punchthrough; i.e., there are no dopant ion implantations, so our model's dependence on a constant body doping density is especially suitable. These well densities were determined by (approximately) matching simulated and measured subthreshold current slopes ($d(\log|I_{DS}|)/d|V_{GS}|$) for $|V_{DS}| = 0.05\text{ V}$ and $V_{BS} = 0\text{ V}$. Using these densities ($1.1 \times 10^{17}\text{ cm}^{-3}$ for the n-well and $1.4 \times 10^{17}\text{ cm}^{-3}$ for the p-well) the program SUMM and the routine MODCHK in MMSPICE provided evaluations of several other key model parameters, including the gate-body work-function difference, the flatband voltage, and the low-field carrier channel mobility. This latter parameter, which was evaluated in SUMM using a measurement-based mobility-doping density relation, was adjusted slightly (10%) to improve the simulations, reflecting the inevitable need for experimental-based tuning of this parameter's evaluation. Parasitic source and drain resistances along with the mobility degradation coefficient (THETA) were estimated from $I_{DS}(V_{DS})$ measurements (for $A_{\text{source}} = A_{\text{drain}} = 40\text{ }\mu\text{m}^2$). Typical values for the LDD/LDS junction

depths, doping densities, and lengths were used, because the actual values were not available to us. The LDD and LDS however are not critical to the comparisons presented in this chapter, and in essence simply provide additional source and drain series resistances.

Figure 5.1 shows $I_{SD}(V_{SG})$ characteristics ($V_{SD} = 0.05, 5$ V; $V_{SB} = 0$ V) for a p-channel MOSFET, and Fig. 5.2 shows $I_{DS}(V_{GS})$ ($V_{DS} = 0.05, 5$ V; $V_{BS} = 0$ V) for a n-channel MOSFET. Notice that the simulations and measurements compare quite well, and despite apparent discrepancies, do provide verification of our model's accuracy. Some discrepancies exist in the currents of the moderate surface-inversion regime, which are approximately the regions $V_{SG} = 0.4$ to 1 V, for p-channel, and $V_{GS} = 0.5$ to 1.1 V, for n-channel; we attribute these to the inherent inaccuracy of the cubic spline.

Figure 5.3 shows measured and simulated $I_{SD}(V_{SD})$ characteristics ($V_{SG} = 1$ to 5 V) for the p-MOSFET, and Fig. 5.4 shows the $I_{DS}(V_{DS})$ characteristics ($V_{GS} = 1$ to 5 V) for the n-MOSFET. These simulations and measurements also compare favorably. Notice these curves coincide at low biases, but diverge at higher biases. We attribute these discrepancies primarily to the lattice-heating effect due to local power dissipation. Additionally, the device manufacturer suggests some discrepancy may be due to the formation of a depletion-charge layer in the polysilicon gate, which diminishes the gate-voltage drive on the transistor. The device self-heating

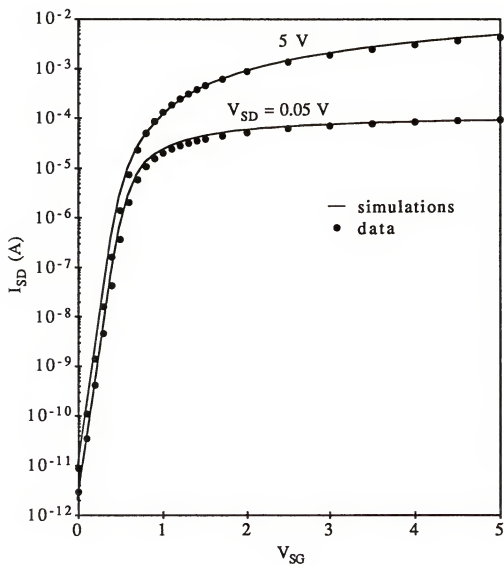


Fig. 5.1. Measured and simulated $I_{SD}(V_{SG})$ characteristics ($V_{SD} = 0.05, 5$ V and $V_{SB} = 0.0$ V) for the p-channel MOSFET.

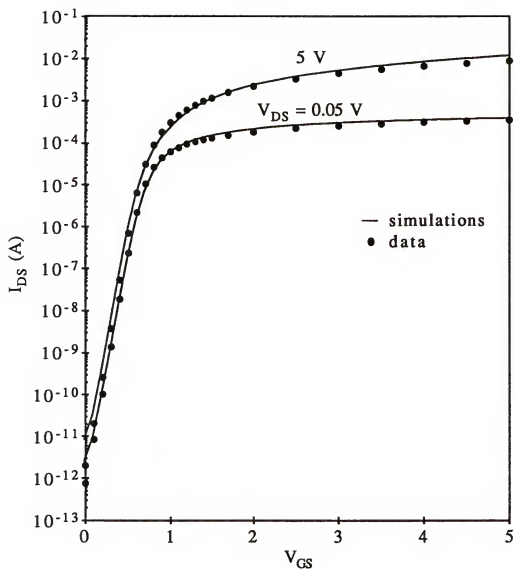


Fig. 5.2. Measured and simulated $I_{DS}(V_{GS})$ characteristics ($V_{DS} = 0.05, 5$ V and $V_{BS} = 0.0$ V) for the n-channel MOSFET.

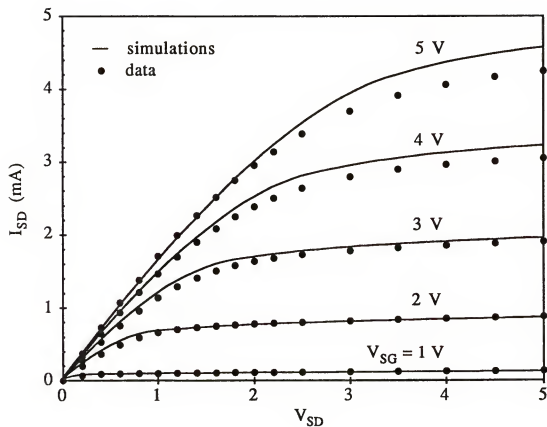


Fig. 5.3. Measured and simulated $I_{SD}(V_{SD})$ characteristics ($V_{SG} = 1$ to 5 V) for the p-channel MOSFET.

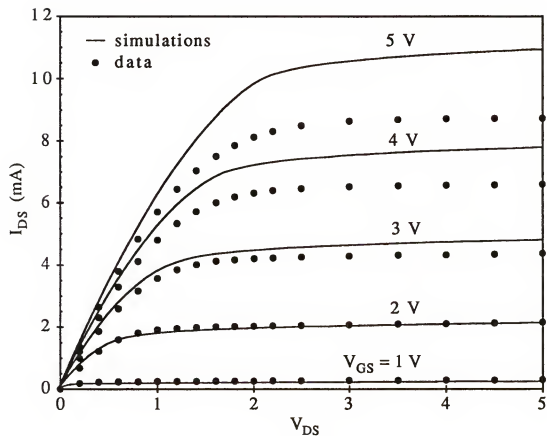


Fig. 5.4. Measured and simulated $I_{DS}(V_{DS})$ characteristics ($V_{GS} = 1$ to 5 V) for the n-channel MOSFET.

causes a reduction in carrier mobility [Sze81], which also lowers channel current. The effect tends to offset channel-length modulation. For these devices lattice heating is quite severe, as indicated by the near-zero slopes of the measured $I_{DS}(V_{DS})$ saturation curves ($V_{GS} = 3$ to 5 V) in Fig. 5.4, and the negative slopes of measurements at higher V_{GS} ($= 6, 7$ V - not shown). Therefore channel-length modulation is not evident in those curves. Note though that our channel-length modulation model is verified at lower gate biases.

It is important to recognize that self-heating can be erroneously accounted for in simulations because of its impact on parameter evaluation. For example, the effect may not be as significant in transient simulations because of the finite time required for self-heating to impact device operation. Therefore, parameter values extracted from measured characteristics such as those in Fig. 5.4 via optimization-based extraction methods (i.e., curve-fitting) could produce erroneous simulations. This suggests an advantage of parameter evaluation methodologies based primarily on structural analyses such as ours.

Comparisons between CMOS digital-signal (transient) circuit measurements and simulations for this technology serve as verification of the new device/circuit simulator. The circuit of interest is shown in Fig. 5.5. The MOSFET widths of each inverter were selected based on the ratio $W_{PMOS}/W_{NMOS} = 2$ (for symmetric inverter signals) for two fan-out (FO) scenarios. For FO = 1 all three inverters are identical

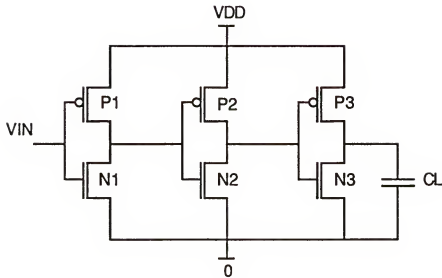


Fig. 5.5. The three-stage CMOS inverter used to verify the mixed device/circuit simulator. Device widths and capacitance C_L values were selected to provide fan-out-of-one and fan-out-of-three scenarios. Switching delays are measured across the second stage.

($W_{\text{NMOS}} = 9 \mu\text{m}$), and the value of the capacitor on the output of the third stage (C_L) is equal to the inverter input capacitance. For $FO = 3$, the inverter input capacitances increase successively by a factor of three ($W_{N2} = 9 \mu\text{m}$), and C_L equals three times the input capacitance of the third stage. For each fan-out scenario, rise-times and fall-times were measured across the second stage based on the times required for the input and output voltages of this stage to reach $V_{DD}/2$; they were then averaged to compute the CMOS inverter switching delays. For $V_{DD} = 5 \text{ V}$, these delays were measured to be 150 ps for $FO = 1$, and 300 ps for $FO = 3$. (We note that for these measurements the circuit configuration for the $FO = 3$ scenario may differ from the one presented above. For example, fan-out of three may have been achieved by driving multiple (parallel) CMOS inverters; as a result, device and circuit parasitics would be different than those of the configuration assumed here, introducing discrepancy in these delay comparisons.)

For the circuit simulations we used the same model parameters as the previous device simulations with extrinsic source and drain resistances (viz., model parameters R_S and R_D) adjusted to account for the various areas of those regions in the circuit. Gate-source and gate-drain overlap capacitances (per μm) were approximated by $C_{ox}DL$, where (parameter) DL is the channel-length reduction due to source/drain diffusion under the gate ($= 0.1 \mu\text{m}$ for this technology). MMSPICE predicted 130 ps for $FO = 1$ and 235 ps

for $FO = 3$. These values are slightly less than the measured values (approximately 13 % for $FO = 1$ and 22 % for $FO = 3$), which we attribute primarily to our uncertainty about the device structure. For example, due to such uncertainties we did not account for gate-source and gate-drain fringing-field capacitances and gate-body overlap capacitances in our simulations; such capacitances increase inverter delay times. Given these conditions our simulations are acceptably accurate estimations of the actual circuit measurements.

5.3 Application to Technology Scaling

In this section we demonstrate the utility of our physical device/circuit simulator in device design. In this exercise we scale the aforementioned $0.8\text{ }\mu\text{m}$ -CMOS technology in accordance with concurrent scaling efforts of the device manufacturer. Due to proprietary considerations our knowledge of the manufacturer's scaled technology is limited, which in fact restricts our device design; for example, we retain the same LDD/LDS junction depth and doping density as the larger technology. The scaling objective is to design a CMOS technology with minimum mask dimension of $0.6\text{ }\mu\text{m}$ for operation at a supply voltage (V_{DD}) of 3.3 V (down from 5 V). Our design criteria include obtaining approximately the same values of threshold voltages as the $0.8\text{ }\mu\text{m}$ -technology, so the MOSFETs will exhibit low subthreshold (leakage) current. We measure threshold voltage as the $|V_{GS}|$ required for $|I_{DS}| = 0.1\text{ (W/L) }\mu\text{A}$, with $V_{BS} = 0\text{ V}$ and $|V_{DS}| = 0.05\text{ V}$. We also

design the scaled devices with the intention of increasing the saturation current (per μm , at $|V_{GS}| = |V_{DS}| = V_{DD}$) above that of the larger technology in order to improve circuit operating speed. Note that our predictions of these currents are dependent on the values of series resistances used in the simulations. From the device manufacturer we learned that the source and drain areas were scaled by a factor of 1.4, which we used to estimate new source and drain parasitic resistances.

We first focus on the n-channel MOSFETs for which the targeted effective channel length is $0.4 \mu\text{m}$. For this exercise we consider as technology parameters only the body doping density and gate oxide thickness. The designed reductions in effective channel length and supply voltage define a scaling factor for these parameters of 1.5 based on the crude "constant field" scaling rules [Den74], which in fact do not account for short-channel effects. Regardless, from these rules we made initial estimates of the new (constant) p-well doping density ($2.1 \times 10^{17} \text{ cm}^{-3}$) and gate oxide thickness (10 nm). SUMM/MMSPICE simulations predicted a n-MOSFET threshold reduction of more than 0.1 V due to scaling and short-channel effects. Therefore, we raised the p-well doping density to increase the threshold voltage. Increasing this density also had the effect of lowering saturation current via channel mobility. (Channel mobility was evaluated by SUMM then tuned by 10%, the same factor used in the $0.8 \mu\text{m}$ -device comparisons of the previous section.) To

raise the saturation current we reduced the gate oxide thickness. In accord with the manufacturer's target thickness we settled on $t_{ox} = 8$ nm. However, reducing t_{ox} also reduced the threshold voltage. This indicates an important design tradeoff between threshold voltage and saturation current that stresses the need for physical modeling. Because our model relies minimally on empiricism, the correlations between threshold voltage and saturation current (as functions of model parameters such as L , NB , UO and TOX) are accounted for physically and accurately. We settled on a final p-well doping density of $3.5 \times 10^{17} \text{ cm}^{-3}$. Figure 5.6 compares predicted $I_{DS}(V_{GS})$ characteristics for this device with the simulations of the $0.8 \text{ } \mu\text{m}$ n-MOSFET, which are also shown in Fig. 5.2; for both MOSFETs the effective channel width is $19 \text{ } \mu\text{m}$. The scaled n-MOSFET threshold voltage is 0.6 V , compared with 0.65 V of the larger MOSFET, and the drain-source current density at $V_{GS} = V_{DS} = V_{DD}$ is $0.46 \text{ mA}/\mu\text{m}$ for the scaled device compared with $0.57 \text{ mA}/\mu\text{m}$ for the $0.8 \text{ } \mu\text{m}$ MOSFET. The values for the scaled n-MOSFET almost precisely match those of the manufacturer's design. Notice the current density for the scaled device is less than that for the simulated $0.8 \text{ } \mu\text{m}$ n-MOSFET, a result of having to use high p-well doping density to avoid undesirable short-channel effects in our simulations. (Further reductions in oxide thickness to increase current density may be limited by manufacturing capability.)

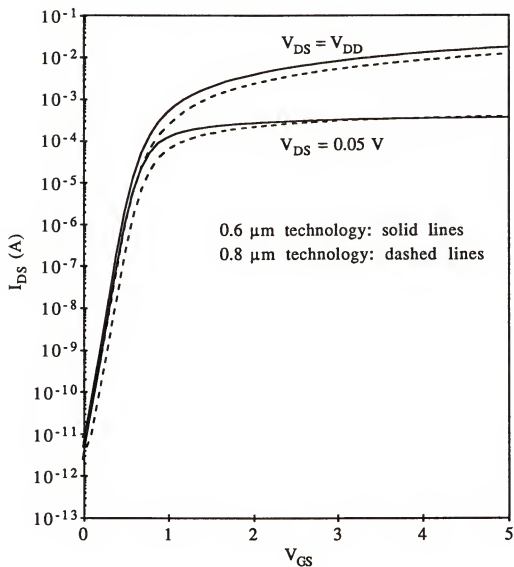


Fig. 5.6. Simulated $I_{DS}(V_{GS})$ characteristics ($V_{DS} = 0.05$ V, V_{DD} and $V_{BS} = 0.0$ V) for the n-channel MOSFETs of the 0.8 μm technology ($V_{DD} = 5$ V) and scaled 0.6 μm technology ($V_{DD} = 3.3$ V).

The scaled p-channel MOSFETs are designed to have effective channel lengths of $0.35\ \mu\text{m}$ (set by the manufacturer) and gate oxide thickness of $8\ \text{nm}$, as established in the n-MOSFET scaling. Therefore the only p-MOSFET design parameter used here is the (constant) n-well doping density. Based on the tradeoff between low subthreshold current (high threshold voltage) and high saturation current, we settled on an n-well doping density of $2.7 \times 10^{17}\ \text{cm}^{-3}$. Predicted $I_{SD}(V_{SG})$ characteristics for the scaled p-channel MOSFET are compared in Fig. 5.7 with simulations of the $0.8\ \mu\text{m}$ -device. Here the threshold voltage has reduced from $0.6\ \text{V}$ to $0.5\ \text{V}$, and the saturation current density at $|V_{GS}| = |V_{DS}| = V_{DD}$ has diminished from $0.24\ \text{mA}/\mu\text{m}$ to $0.17\ \text{mA}/\mu\text{m}$. Manufacturer's measurements of these (preliminary) devices yielded a threshold voltage of $0.55\ \text{V}$ and saturation current density of $0.21\ \text{mA}/\mu\text{m}$. We attribute much of this discrepancy to our uncertainties about the device structure, e.g., with regard to series resistances, as well as uncertainty in the process-dependent channel mobility.

The resulting switching delays characterizing the scaled $0.6\ \mu\text{m}$ technology were derived from transient simulations of the circuit in Fig. 5.5 with $V_{DD} = 3.3\ \text{V}$, using the same widths as those of the $0.8\ \mu\text{m}$ -technology simulations. They were predicted to be $155\ \text{ps}$ ($FO = 1$) and $285\ \text{ps}$ ($FO = 3$), implying no speed enhancement; these delays are in fact somewhat longer than those predicted for the $0.8\ \mu\text{m}$

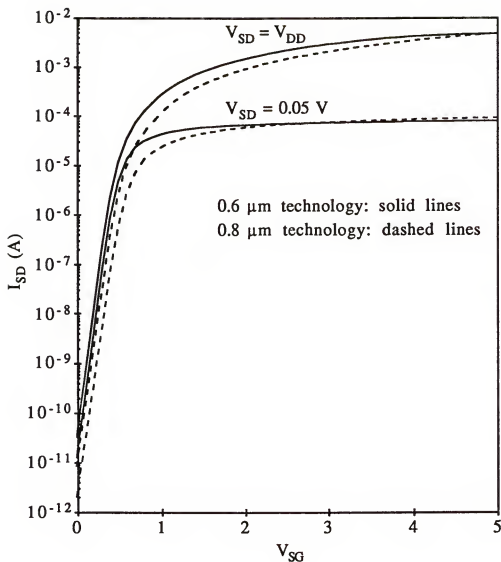


Fig. 5.7. Simulated $I_{SD}(V_{SG})$ characteristics ($V_{SD} = 0.05$ V, V_{DD} and $V_{SD} = 0.0$ V) for the p-channel MOSFETs of the 0.8 μm technology ($V_{DD} = 5$ V) and scaled 0.6 μm technology ($V_{DD} = 3.3$ V).

technology ($V_{DD} = 5$ V). The increases are consistent with the noted reductions in drive current that resulted from the technology scaling.

CHAPTER 6

SUMMARY AND SUGGESTIONS FOR FUTURE WORK

In this dissertation a new approach to integrated process/device/circuit simulation was presented that relies on pragmatic seminumerical physical device models. A short-channel MOSFET model for this application was derived and implemented into the circuit simulator SPICE2 to create MMSPICE-3. The model relies on quasi-two-dimensional analyses of the MOSFET structure to enable predictive coupled device and circuit simulations. The parameter evaluator SUMM was extended to include evaluation algorithms for several key MOSFET model parameters based on analyses of the device structure, forming the link between MMSPICE-3 and a process simulator. Because the parameters of our model are structurally based, their evaluation is straightforward, and the model readily provides insight regarding physical device mechanisms. This is in contrast to empirical models, e.g., BSIM [She87], which rely heavily on "fitting" parameters that tend to obscure the device physics. PISCES simulations and device measurements provided support for SUMM/MMSPICE-3 and the pragmatic tool-integration approach. The utility of the tool was demonstrated in an application to optimal scaling of a BiCMOS technology. Both SUMM and MMSPICE-3 are available at the University of Florida

The following tasks are suggested to enhance the integrated SUMM/MMSPICE-3 simulator that is now capable of bipolar, CMOS, and BiCMOS mixed-mode device/circuit simulation and technology CAD.

(1) In Chapter 5 comparisons between measurements and model simulations suggested several tasks for improving this model and all circuit-based MOSFET models. Perhaps the most important among those is the development of a model for lattice self-heating due to high local power dissipation. Continued measurement-based verification of the MOSFET model should be undertaken to help identify other model features that may benefit from refinement.

(2) Model applications described in Chapters 2 and 5 imply the inevitable need for experimentally based tuning of the model parameter evaluations. Such tuning could be accomplished by incorporating new routines in SUMM that would automatically calibrate the parameter evaluation algorithms using user-supplied device measurements.

(3) Extending the capabilities of SUMM to include evaluation of MOSFET parameters from two-dimensional structural information may be useful. This would enable straightforward evaluation of junction depths and lateral structure parameters; and the evaluation analyses of other parameters, e.g., the effective body doping density, could include accounting for two-dimensional body-doping nonuniformities. Such nonuniformities can result from lateral body-dopant

diffusion that is enhanced by drain and source implant damage.

(4) To improve the robustness of the simulator, the model and SUMM could be modified to account for buried-channel MOSFETs which are part of some CMOS technologies.

(5) Inclusion of "non-local heating" effects [Ant91] would provide a useful enhancement of the MOSFET model, e.g., for addressing hot-carrier reliability, for simulation of deep-submicron (e.g., $L < 0.25 \mu\text{m}$) devices. The characterization of the impact-ionization current should be upgraded to account for non-local effects on $(M - 1)$ in highly scaled devices.

APPENDIX
EVALUATION OF THE MMSPICE BJT
MODEL [Jeo89] PARAMETERS JE0 AND TE

For an npn transistor, the parameter TE, the minority-carrier (hole) charge-control time for the quasi-neutral emitter (QNE), relates the hole charge (Q_p) in the silicon/polysilicon emitter to the hole current injected into the emitter, I_{pE} . This current has a saturation current density represented by the model parameter JE0. This appendix presents evaluation algorithms for these parameters that are revisions of those in the original SUMM [Gre90]. The evaluation of TE and JE0 is based on a physical analytic characterization [Fos81] of the minority-hole transport which accounts for bandgap-narrowing and Auger-recombination processes in the n^+ emitter. The parameter JE0 is evaluated analytically as [Fos81]

$$JE0 = \frac{qn_i^2}{N_{D(eff)}} \left(\frac{1}{S_p} + \frac{W_H}{D_{p0}} \right)^{-1} \left(1 + \frac{\tau_t}{\tau_A} \right), \quad (A1)$$

where W_H is the width of the QNE region of doping density at least equal to an effective donor density $N_{D(eff)}$. Because of bandgap narrowing and electron degeneracy, $N_{D(eff)} \cong 10^{18} \text{ cm}^{-3} < N_D$, where N_D is the spatially-dependent QNE doping density. In (A1), D_{p0} is the hole diffusivity at the silicon surface,

which is computed from a carrier mobility-doping density dependence; the quantity

$$\bar{\tau}_A \Delta \frac{1}{C_p N_D^2 (W_H/2)} \quad (A2)$$

is an average Auger lifetime for holes in the QNE; and τ_t is the hole transit time across the QNE, expressed as

$$\tau_t = \frac{W_H^2}{2D_{p0}} + \frac{W_H}{S_p} . \quad (A3)$$

Because of the complex nature of the polysilicon emitter contact [Yun88], the effective hole surface recombination velocity, S_p , cannot be accurately determined theoretically; therefore it must be determined by the user by some other (experimental) means.

The relationship between TE and JEO is defined as

$$Q_{p(\text{total})} = Q_{p(\text{QNE})} + Q_{p(\text{poly})} \equiv TE * I_{pE} , \quad (A4)$$

where

$$Q_{p(\text{QNE})} = \left(\frac{1}{\tau_t} + \frac{1}{\tau_A} \right)^{-1} I_{pE} \quad (A5)$$

and

$$Q_{p(\text{poly})} \equiv qA p_{\text{poly}} L_{p(\text{poly})} = \frac{L_{p(\text{poly})}}{S_p} I_{pE} . \quad (\text{A6})$$

In (A6), $L_{p(\text{poly})}^2 \equiv D_{p(\text{poly})} \tau_{A(\text{poly})}$, $D_{p(\text{poly})} \equiv D_{p0}$, $\tau_{A(\text{poly})} \equiv (C_A N_{DS}^2)^{-1}$, and p_{poly} is the hole density at the polysilicon contact. The doping density in the polysilicon is assumed to be uniform and equal to that (N_{DS}) at the surface of the underlying silicon emitter. Equation (A6) derives under the assumptions that the polysilicon thickness is greater than the hole diffusion length $L_{p(\text{poly})}$, and that

$$I_{pE} = qAS_p p_{\text{poly}} \equiv qA \frac{D_{p(\text{poly})}}{L_{p(\text{poly})}} p_{\text{poly}} , \quad (\text{A7})$$

which reflects our assumption of an extended-emitter structure [Yun88], with p_{poly} equal to the hole density at the silicon-emitter surface. Because of the assumption of the extended emitter, (A6) is not applicable to the case of a polysilicon contact with an interfacial oxide. Such an oxide would suppress carrier injection, and therefore reduce S_p and $Q_{p(\text{poly})}$. Obviously (A6) can predict an unbounded increase in $Q_{p(\text{poly})}$ with decreasing S_p . To avoid this inconsistency, we define a minimum S_p for which (A6) is valid; based on experimental data [Yun88], we choose 10^4 cm/s as this minimum. Then for $S_p < 10^4$ cm/s, we assume $Q_{p(\text{poly})} = 0$; hence (A4) and (A5) define TE. For $S_p > 10^4$ cm/s though, (A6) and (A7) give

$$Q_{p(\text{poly})} = \frac{D_{p(\text{poly})}}{S_p^2} I_{pE} , \quad (\text{A8})$$

and (A4) yields

$$TE = \left(\frac{1}{\tau_t} + \frac{1}{\tau_A} \right)^{-1} + \frac{D_{p(\text{poly})}}{S_p^2} . \quad (\text{A9})$$

REFERENCES

- Ant88 P. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE*, McGraw-Hill, New York, 1988.
- Ant91 D. A. Antoniadis and J. E. Chung, "Physics and Technology of Ultra Short-Channel MOSFET Devices", IEEE IEDM-91, Dig. Tech. Papers, p. 21, 1991.
- Cha87 P. C. Chan, R. Liu, S. K. Lau, and M. Pinto-Guedes, "A Subthreshold Conduction Model for Circuit Simulation of Submicron MOSFET," IEEE Trans. on Computer-Aided Des., vol. 6,, p. 574, 1987.
- Cho91 J. Y. Choi, "Modeling and Simulation of the Fully Depleted Silicon-on-Insulator MOSFET for Submicron CMOS IC Design," Ph.D. dissertation, University of Florida, Gainesville, 1991.
- Den74 R. H. Dennard, F. H. Gaensslen, H. N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of Ion-Implanted MOSFETs with very Small Physical Dimensions," IEEE J. Solid-State Circuits, vol. SC-9, p. 256, 1974.
- Dor92 M. J. van Dort, P. H. Woerlee, A. J. Walker, C. A. H. Juffermans, and H. Lifka, "Influence of High Substrate Doping Levels on the Threshold Voltage and the Mobility of Deep-Submicrometer MOSFET's," IEEE Trans. Electron Devices, vol. ED-39,, p. 932, 1992.
- Eva90 I. R. Evans, N. Morris, P. J. Mole, S. C. Quinlan, I. Jackson, and P. A. Murkin, "Optimization of Polysilicon Emitters for BiCMOS Transistor Design," IEEE Trans. Electron Devices, vol. ED-37, p. 2343, 1990.
- Fos81 J. G. Fossum and M. A. Shibib, "An Analytic Model for Minority-Carrier Transport in Heavily Doped Regions of Silicon Devices," IEEE Trans. on Electron Devices, vol. ED-28, p. 1018, 1981.
- Get78 I. E. Getreu, *Modeling the Bipolar Transistor*, New York, Elsevier, 1978.

- Gre90 K. R. Green, "SUMM: A SUPREM-3/MMSPICE-1 Integrator for Bipolar Technology CAD," M.S. thesis, University of Florida, Gainesville, 1990.
- Jeo89 H. Jeong and J. G. Fossum, "A Charge-Based Large-Signal Bipolar Transistor Model for Device and Circuit Simulation," IEEE Trans. Electron Devices, vol. ED-36, p. 124, 1989.
- Jeo90 H. Jeong, J. G. Fossum, and D. K. FitzPatrick, "MMSPICE: A Semi-Numerical Mixed-Mode Device/Circuit Simulator for Advanced Bipolar Technology CAD," Solid-State Electron., vol. 33, p. 1283, 1990.
- Law90 M. Law, E. Solley, D. Burk, M. Liang, and D. Apte, "Improved Simulation of Advanced Bipolar Structures," Ext. Abs. Semiconductor Research Corp. TECHCON '90 Conf. (San Jose, CA), p. 65, 1990.
- Llo90 P. Lloyd, H. K. Dirks, E. J. Prendergast, and K. Singhal, "Technology CAD for Competitive Products," IEEE Trans. Computer-Aided Design, vol. 9, p. 1209, 1990.
- Man77 Y. A. El-Mansy and A. R. Boothroyd, "A Simple Two-Dimensional Model for IGFET Operation in the Saturation Region," IEEE Trans. Electron Devices, vol ED-24, p. 254, 1977.
- Mar88 J. Mar, "CAD for Manufacturing: Why Its Time Is Now," Ext. Abs. for Semiconductor Research Corporation TECHCON '88 Conference (Dallas, TX), p. 277, 1988.
- Mul86 R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, Wiley, New York, 1986.
- Nag75 L. W. Nagel, "SPICE2: A Computer Program to Simulate Semiconductor Circuits," Electron. Res. Lab. Memo. ERL-M520, University of California, Berkeley, 1975.
- Oh80 S.-Y. Oh, D. E. Ward, and R. W. Dutton, "Transient Analysis of MOS Transistors," IEEE Trans. Electron Devices, vol. ED-27, p. 1571, 1980.
- Pin84 M. R. Pinto, C. Rafferty, and R. W. Dutton, "PISCES II: Poison and Continuity Equation Solver," Stanford Electronics Lab., Stanford University, Tech. Rep., 1984.

- Pre89 W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, 1989.
- She87 B. J. Sheu, D. L. Scharfetter, P. K. Ko, and M. C. Jeng, "BSIM: Berkeley Short-Channel IGFET Model for MOS Transistors," IEEE J. Solid-State Circuits, vol. SC-22, p. 558, 1987.
- Slo87 J. W. Slotboom, G. Streutker, G. J. T. Davis and P. B. Hartog, "Surface Impact Ionization in Silicon Devices," IEEE IEDM-87, Dig. Tech. Papers, p. 494, 1987.
- Sze81 S. M. Sze, *Physics of Semiconductor Devices*, 2nd Ed., Wiley, New York, NY, 1981.
- Tay78 G. W. Taylor, "Subthreshold Conduction in MOSFET's," IEEE Trans. Electron Devices, vol. ED-25, p. 337, 1978.
- Tec88 Technology Modeling Associates, Inc., *TMA-SUPREM-3: One-Dimensional Process Analysis Program*, Palo Alto, CA, 1988.
- Tec89 Technology Modeling Associates, Inc., *TMA-PISCES-2B: Two-Dimensional Device Analysis Program*, Palo Alto, CA, 1989.
- Tro79 R. R. Troutman, "VLSI Limitations from Drain-Induced Barrier Lowering," IEEE Trans. Electron Devices, vol. ED-26, p. 461, 1979.
- Tsi82 Y. Tsiividis, "Moderate Inversion in MOS Devices," Solid-State Electron., vol. 25, p. 1099, 1982.
- Van90 P. Vande Voorde, D. Pettengill, and S.-Y. Oh, "Hybrid Simulation and Sensitivity Analysis for Advanced Bipolar Device Design and Process Development," Ext. Abs. IEEE 1990 BCTM, p. 114, 1990.
- Vee88 S. Veeraraghavan and J. G. Fossum, "A Physical Short-Channel Model for the Thin-Film SOI MOSFET Applicable to Device and Circuit CAD," IEEE Trans. Electron Devices, vol. ED-35, p. 1866, 1988.
- Yau74 L. D. Yau, "A Simple Theory to Predict the Threshold Voltage of Short-Channel IGFET's," Solid-State Electron., vol. 17, p. 1059, 1974.

- Yun88 S.-Y. Yung and D. E. Burk, "Informed Device Design and Gain-Speed Trade-Off for Self-Aligned Polysilicon-Emitter Transistors," Solid-State Electron., vol. 31, p. 1139, 1988.
- Zde87 P. J. Zdebel, R. J. Balda, B.-Y. Hwang, V. d. l. Torre, and A. Wagner, "MOSAIC III - A High Performance Bipolar Technology with Advanced Self-Aligned Devices," Proc. IEEE BCTM, p. 172, 1987.

BIOGRAPHICAL SKETCH

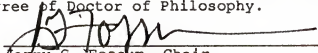
Keith R. Green was born in Wilmington, Delaware, in 1966. He received the B.S. degree in electrical engineering in 1988 from the University of Delaware, Newark. In 1990 he received the M.S. degree in electrical engineering from the University of Florida, Gainesville, and in the same year began working towards the Ph.D. degree in electrical engineering.

He received the 1990 Outstanding Master's Thesis Award from the Department of Electrical Engineering, University of Florida.

He spent the summers of 1989 and 1990 working at Harris Semiconductor and National Semiconductor. Upon graduation he will join Texas Instruments, Inc. in Dallas, Texas, as a Mixed Signal IC Designer working in areas of cryogenic-MOSFET modeling and technology characterization.

He is a member of IEEE and Tau Beta Pi.

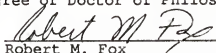
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Jerry G. Fossum, Chair
Professor of Electrical Engineering


I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Dorothea E. Burk
Professor of Electrical Engineering

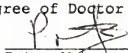
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Robert M. Fox
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Mark E. Law
Associate Professor of Electrical Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.


Peter Mataga
Assistant Professor of Aerospace Engineering, Mechanics and Engineering Science

This dissertation was submitted to the Graduate Faculty of the College of Engineering and the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1993



Winfred M. Phillips
Dean, College of Engineering

Dean, Graduate School